PCT/GB 2003 / 002244

Rec'd PCT/PTO 23 NOV 2004

10/515985

# PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ



I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

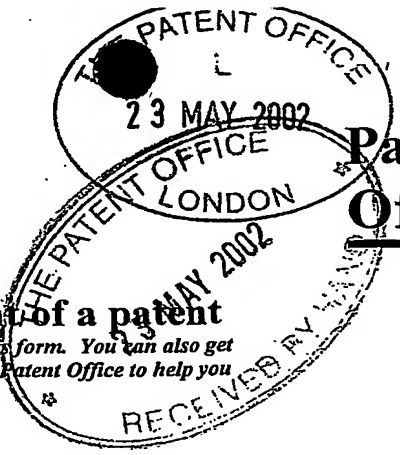
Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed R. Mahoney

Dated 16 JUN 2003

Patents Form 1/77

Patents Act 1977
(1990)




The Patent Office

24MAY02 E721005-1 D01038
P01/7700 0.00-0211943.6

# Request for grant of a patent

*(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)*

The Patent Office
Cardiff Road
Newport
South Wales
NP10 8QQ

1. Your reference P706487GB/DE/47491

2. Patent application number *(The Patent Office will fill in this part)* 0211943.6

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*

Sheffield Hallam University
Research and Business Development Department
City Campus
Howard Street
Sheffield
S1 1WB
United Kingdom

Patents ADP number *(if you know it)* 74441100 03

If the applicant is a corporate body, give the country/state of its incorporation: England

4. Title of the invention: Anti-Helicobacter Activity of Celery Seed Components

5. Name of your agent *(if you have one)*
"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

Dr David Elsy
WITHERS & ROGERS
Goldings House
2 Hays Lane
London
SE1 2HW

Patents ADP number *(if you know it)* 1776001

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or each of these earlier applications and *(if you know it)* the or each application number

| Country | Priority application number *(if you know it)* | Date of filing *(day / month / year)* |
|---|---|---|

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

| Number of earlier application | Date of filing *(day / month / year)* |
|---|---|

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(answer 'Yes' if:*
*a) any applicant named in part 3 is not an inventor, or*
*b) there is an inventor who is not named as an applicant, or*
*c) any named applicant is a corporate body.*
*See note (d))*

YES

**Patents Form 1/77**

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

| | |
|---|---|
| Continuation sheets of this form | |
| Description | 14 |
| Claim(s) | DMC |
| Abstract | |
| Drawing (s) | 12+12 |

10. If you are also filing any of the following, state how many against each item.

| | |
|---|---|
| Priority documents | |
| Translations of priority documents | |
| Statement of inventorship and right to grant of a patent *(Patents Form 7/77)* | |
| Request for preliminary examination and search *(Patents Form 9/77)* | |
| Request for substantive examination *(Patents Form 10/77)* | |
| Any other documents *(please specify)* | |

11. I/We request the grant of a patent on the basis of this application.

Signature Withoulagers Date 22 May 2002

12. Name and daytime telephone number of person to contact in the United Kingdom David Elsy 01926 336111

**Warning**

*After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least six weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.*

**Notes**

a) *If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500 505.*
b) *Write your answers in capital letters using black ink or you may type them.*
c) *If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.*
d) *If you have answered 'Yes' Patents Form 7/77 will need to be filed.*
e) *Once you have filled in the form you must remember to sign and date it.*
f) *For details of the fee and ways to pay please contact the Patent Office.*

# Anti-Helicobacter Activity of Celery Seed Components

The invention relates to the use of biologically active celery seed extracts to inhibit the growth and replication of the bacterium, *Helicobacter pylori.*

Arthritis and rheumatism are important world-wide problems. Around 1% of the UK population are affected at some stage in life. Complaints of this nature not only cause significant disability but may also have a severely detrimental effect on the psychological state of the sufferers. Conventionally these complaints are treated with analgesic/antipyretic drugs and non-steroidal anti-inflammatory drugs (NSAIDs). However NSAIDs can have serious side effects, such as gastrotoxicity, causing for example gastric ulceration, and hence research has been made into alternative sources of anti-inflammatory drugs. In particular compounds extracted from higher plants have been considered. Lewis *et al* (1985) and Whitehouse *et al* (1999) found that the extracts of celery (*Apium graveolens*) (CSE) had significant anti-inflammatory activty in animal models with reduced adverse effects. A further risk factor in the pathogenesis of peptic ulcer disease is *H.pylori* infection. Chan (1997) found that eradication of *H.pylori* before NSAID therapy reduced the risk of ulcer development by about fourfold. PCT/US99/25873 discloses the use of celery seed extract for the prevention and treatment of pain, inflammation and gastrointestinal irritation.

The inventors have surprisingly found that components of celery seed extract may be used to control the growth of *Helicobacter pylori.*

The invention provides the use of celery seed or celery seed extract (CSE) for the inhibition of growth and replication of *Helicobacter pylori.*

A preferred CSE is produced by supercritical fluid extraction of the starting product. By CSE we mean a natural product derived from celery seed, or a pharmaceutical equivalent thereof. This is preferably an ethanol/water extract, especially 50% to 90%,

60% to 85%, most preferably an 80% Vol:Vol ethanol/water extract. The term includes the isolated compounds obtainable from CSE.

Preferably the active component of the celery seed extract is selected from the group: 3-n-butyl 4,5-dihydrolphthalide, 3-n-butyl phthalide, $\alpha$-Eudesmol, $\beta$-Eudesmol dioctyl phthalate and cis, cis-9,12-Octadecadienoic acid.

The invention further provides a pharmaceutical composition for the inhibition of growth and replication of *Helicobacter pylori*, comprising celery seed extract.

Also provided is the use of celery seed extract in the preparation of a pharmaceutical composition for the treatment of *Helicobacter pylori* infection.

Preferably the *H.pylori* infection is in a mammal, such as a human. Preferably the infection is within the digestive tract, especially the stomach of the mammal.

The pharmaceutical composition may be administered orally, e.g. in the form of an oral suspension. solution or tablet. Dosages may be 300-2000 mg. daily in divided doses preferably or even higher.

The pharmaceutical composition may comprise one or more pharmaceutically acceptable carriers, bulking agents or excipients known in the art (e.g. in the form of a tablet or injectable solution).

A further aspect of the invention provides celery seed or celery seed extract for use in the manufacture of a medicament to treat a *Helicobacter pylori* infection.

The invention will now be described in detail with reference to the figures in which:

Table 1 shows the effect of the crude extract of CSE on the growth of different strains (3330, 3336 and 3339) of *H.pylori*.

Table 2 shows the distribution of antimicrobial activity against *H. pylori* (strain 3339) in the crude extract and different fractions of CSE.

Table 3 shows antimicrobial activity of the subfractions from pet. ether fraction against *H. pylori* (strain 3339).

Table 4 shows antimicrobial activities of compounds from subfractions 6 and 10 against *H. pylori* (strain 3339).

Fig.1 shows the effect of CSE crude extract on the growth of the strains (3330, 3336, 3339) of *H.pylori*

Fig.2 shows the bioassay-guided fractionation scheme of celery seed extract (antimicrobial agents enclosed in boxes).

Fig.3 shows the antimicrobial activity of pet. ether fraction and subfractions 6 and 10 against *H.pylori* (strain 3339).

Fig.4 shows the analytical separation of mixture from subfraction 10. Column: Nucleosil® C18, 250 x 4.6 mm. I.D.; Mobile phase: ACN/water (60:40); Flow rate: 1.0 m./min; Detection: UV @ 236 nm; Injection volume: 10 $\mu$g in 1 ml of 40% ACN in water ; Temperature: Ambient; ATT:3.

Fig.5 shows the antimicrobial activities of compounds against *H.pylori* (strain 3339)

Fig.6 shows the EI-MS spectrum of compound 6-1

Fig.7 shows the $^{1}H$ NMR spectrum of compound 6-1

Fig.8 shows the $^{13}C$ NMR spectrum of compound 6-1

Fig.9 shows the EI-MS spectrum of compiund 6-1

Fig.10 shows the EI-MS spectrum of compound 6-3

Fig.11 shows the EI-MS spectrum of compound 6-4

Fig.12 shows theEI-MS spectrum of compound 10-1

**Antimicrobial test**

Bacterial strains

Three strains of *H.pylori* (3330, 3336 and 3339) isolated from British patients with gastric ulcer (duodenal ulcer or gastritis) were studied. The identities of *H. pylori* were confirmed by Gram stain and urease reaction. The bacteria were stored at -80°C in aliquots of 1ml of brocella broth containing 15% (v/v) glycerol (Kitsos and Stadtlander, 1998).

Celery seed extract (CSE)

Test CSE was provided as dark green highly viscous liquid (supplied by Beagle International Pty. Ltd. Nerang, Qld., Australia). Initially CSE was dissolved in dimethylsulfoxide (DMSO) as stock solution (`100mg/ml. final DMSO concentration in cultures ≤1%).

Media

For the Brucella broth (BB), (BBL, USA), Brucella (28g) was added to IL of distilled water. After the medium was autoclaved at 120°C for 15 mins, fetal bovine serum (50 ml) was added (Morgan *et al*, 1987).

Inocula

Thawed isolates were inoculated onto chocolate agar plates (Mërieux) and incubated under microaerophillic conditions (85%$N_2$, 10%$CO_2$, 5%$0_2$) for 48 h at 37°C. Colonies were suspended in 5ml of Brucella broth and adjusted to a turbidity equivalent to a No.2 McFarland standard (approximately $6x10^8$ CFU/ml) for broth dilution method. The final inoculum was $10^7$ CFU/ml for agar dilution method by a further 50-fold dilution.

Broth dilution test

The CSE suspension (1mg/ml) was serially two-fold diluted in BB. The concentrations (1000. 500, 250, 125, and 62.5 µg/ml) were obtained. The solutions were added to the

column wells of 24-well plate each in equal volume (1ml/well). 20µl of cell suspension was inoculated into each row wells of 24-well plates (except last row wells). The culture dishes were gently agitated folllowing the addition of the inoculum and then placed at 37°C under microaerophilic conditions for three days. At the end of incubation, 1ml of bacterial culture solution from each well were diluted to one in a million dilution ($10^{-6}$). Then 20 µl aliquots from each solution were transferred to columbia agars and incubated for an additional three days. Generally, only spots with between 7-11 colonies were counted. Growth was determined on the basis of calculating the number of bacteria per millilitre (numbers of bacteria/ml = numbers of colonies on plate x reciprocal of dilution of sample). Bacteria growth, culture medium and extract controls were run in parallel. (Osato *et al*, 1999).

**Chromatographic Methods**

Column chromatography was performed on silica gel 60 (40-60 µm, Merck). Analytical thin layer chromatography (TLC) was carried out on precoated silica gel 60 $F_{254}$ plates (layer thickness 0.2 mm, Merck), developed with the following solvent, hexane-EtOAc (70:30), chloroform-methanol (98: 2). For isolation monitoring, spots were located by their absorption under ultraviolet (UV) light (254 and 366 nm) directly. After that the plates were sprayed with anisaldehyde reagent and heated at 110°C for 5 min (Dey and Harborne, 1991).

**HPLC (1090 LC, Hewlett Packard, UK) analytical and semi-preparative purification**

Analytical conditions:

Analytical column: Nucelosil® C18, particle size 5µm, 250 x 4.6 mm I.D., catalogue No.89141 (Alltech, Carnforth. Lancashire, UK)

Mobile phase: acetonitrile/water (60:40)

Flow rate: 1.0 ml/min

Injection volume: 10µl

Detection: UV @ 236 nm

Sample: mixture of compounds 10-2, 10-3 and 10-4 (Conc.= 500 µg/ml)

Temperature: ambient

ATT: 3

Semi-preparative conditions:

Semi-preparative column: Luna C18(2), particle size 5µm, 250 x 10.00 mm I.D., catalogue No.00G-4252-NO (Phenomenex, Macclesfield, Cheshire, UK)

Mobile phase: acetonitrile/water (60:40)

Flow rate: 5.0 ml/min

Injection volume: 100µl

Detection: UV @ 236 nm

Sample: mixture of compounds 10-2, 10-3 and 10-4 (Conc. = 5mg/ml)

Temperature: ambient

ATT: 6

**Spectroscopic Methods**

Mass spectrometry (MS)

The Mass spectra were recorded on a VG 70/70 Sector Mass Spectrometer instrument (Micromass, Manchester, UK) in the Laboratory of Biomedical research centre (Sheffield Hallam University).

Nuclear magnetic resonance (NMR)

NMR spectra were recorded in $CDC1_3$ at RT on a Bruker Unity Ac 250 MHz ($^1H$ 250MHz; $^{13}C$, 62.9 Mhz).

## Results and Discussion

The 80% ethanol extract exhibited appreciable antimicrobial activity at the minimum inhibitory concentrations (MIC) of 250, 125 and 125µg/ml, respectively, against *H.*

*pylori* strains 3330, 3336 and 3339. The results of antimicrobial activity of CSE are given in Table 1 and Fig.1. The bioassay-guided fractionation scheme of CSE is illustrated in Fig.2. The fractionation for the isolation of the active compounds was performed from the 80% ethanol extract of CSE. The susceptibility of *H. pylori* strain 3339 was higher than 3330 and 3336. Later, in antimicrobial activity testing of fractions and subfractions of CSE, only *H. pylori* 3339 strain was chosen for fractionation guide. The residue of 80% ethanol extract of CSE was subsequently successively partitioned with organic solvents and water. The activity emerged predominantly in the petroluem ether layer (MIC = 15.625 μg/ml) as compared to the other solvents, diethyl ether (MIC=125μg/ml), ethyl acetate (MIC > 500 μg/ml) and water (MIC > 500 μg/ml) (Table 2).

The petroleum ether fraction was directly subjected to column chromatography on silica gel with hexane, hexane-EtOAc (99:1), hexane-EtOAc (95:5), hexane-EtOAc (70:30) and EtOAc as eluent. Fractions with the same retardation factors were combined to yield 11 major fractions. Each subfraction was tested for antibacterial activity against *H. pylori*. The results of the antimicrobial testing of the different subfractions are shown in Table 3. The most pronounced antimicrobial activity successively resided in the subfraction 6 eluted with hexane-EtOAc (95:5) (MIC = 15.625 μg/ml) and the subfraction 10 eluted with hexane-EtOAc (70:30) (MIC = 15.625 μg/ml (Fig.3). Subfraction 6 was further purified by silica gel column chromatography (hexane-ether, 10:1, as solvent) and preparative TLC using chloroform/pet. ether (3:1) to yield compounds 6-1, 6-2, 6-3 and 6-4. Subfraction 10 was further purified with hexane-ether (7:3) as mobile phase to afford a pure compound 10-1 and a mixture. The mixture was dissolved in 40% ACN in water and passed through the DPA-6S SPE column (Supelco, UK) to remove the chlorophyll. The eluate with methanol was evaporated to dryness and reconstituted in 40% ACN in water for HPLC analysis. It was separated into three compounds 10-2, 10-3 and 10-4 by analytical HPLC using ACN/water (60:40) as mobile phase (Fig.4). Large quantity of individual pure compounds will be obtained by semi-preparative HPLC and sent for MS and NMR spectroscopic analysis.

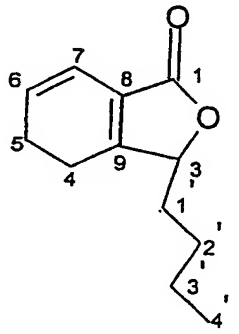
Compounds 6-1, 6-2, 6-3, 10-1 and the combination of 6-1 and 6-3 were evaluated for antimicrobial activity. The results indicated they were partly responsible for the antimicrobial activity of CSE (Table 4 and Fig.5). The mixture of 6-1 and 6-3 by different combination did not exert a syngergistic effect in antimicrobial activity. The mixture of compounds 10-2, 10-3 and 10-4 showed an interesting antimicrobial activity against *H. pylori*. Very recently, Momin and Nair (2001) isolated and characterized three bioactive compounds, sedanolide, senkyunolide-N and senkyunolide-J from CSE with the significant mosquitocidal, nematicidal and antifungal activities. Further study will confirm with MS and NMR data if compounds 10-2, 10-3 and 10-4 are corresponding to sedanolide, senkyunolide-N and sekyunolide-J. The antimicrobial activity of individual compound will be tested as well.

The exact structures are confirmed by comparison of their physical and spectral data ([$\alpha$], 1H and 13NMR) with data in the literature. Structural elucidation of the compounds isolated from active fractions 6 and 10 are given below:

Compound **6-1** was obtained as pale yellow oil with a distinct celery odour. The electron impact mass spectrometry (EI-MS) spectrum (Fig.6) of the compounds showed the molecular ion peak at mass/charge ratio (m/z) 192 (composition, 22.9%), corresponding to the molecular formula $C_{12}H_{16}0_2$. Other major peaks were at m/z (composition, %) 163 (3.6), 135 (5.3), 108 (21.7), 107 (100%), 85 (9.7), 79 (24.3), 77 (24.2) and 57 (14.4).

The $^1$H NMR spectrum (Fig.7) displayed a doublet at 6.12 ppm (1H, J=10 Hz) and a multiplet at 5.9 ppm for the vinyl protons, H-7 ands H-6, respectively, as well as multiplet at 4.9 ppm for H-3. In $^{13}$C NMR spectrum (Fig.8), the signals at 128.4, 116.8 and 124.5 ppm were consistent with disubstituted and tetrasubstituted double bands composed of C-6, C-7 and C-1a, C-3a, respectively. In addition, tetra substituted signals appeared for the side chain (C-1', C-2', C-3', C-4') in the range of 13.8-22.4 ppm. The signals due to C-1, C-4 and C-5 appeared at 161, 31.9 and 26.7 ppm.

On the basis of EI-MS and $^1$H- and $^{13}$C- NMR, compound 6-1 was identified as 3-n-butyl 4,5-dihydrolphthalide (sedanenolide) (Bjeldanes and Kim, 1977).

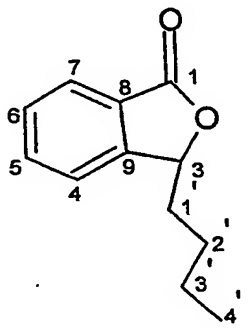



Experimental data

Compound 6-1 EI-MS: m/z 192.3 (calculated for $C_{12}H_{16}O_2$). $^1H$ NMR ($CDCl_3$): $\delta$ 0.9 (t, 3H. J= 7.2, H-4'), 1.2-1.8 [m, 6H, H1(1',2',3')], 2.45 (m, H-4,5), 4.9 (m, 1H, H-3), 5.9 (m. 1H, H-6), 6.2 (d, 1H, J=10, H-7); $^{13}C$ NMR ($CDCl_3$): $\delta$ 13.8-22.4 (C-1', 2', 3', 4'), 26.7-31.8 (C-4.5), 82.5 (C-3), 116.8 (C-7), 128.3 (C-6), 124.5-135 (C-8, 9), 161.4 (C-1).

Compound **6-2** was obtained as pale yellow oil with a distinct celery colour. The EI-MS spectrum (Fig.9) of 6-2 showed the molecular ion peak as mass/charge ratio (m/z) 190, corresponding to the molecular formula $C_{12}H_{14}O_2$. Other major peaks were at m/z 163, 148, 144, 133 (100%), 115, 105, 91 and 77.

On the basis of EI-MS and $^1H$- and $^{13}C$- NMR, compound 6-2 was identified as 3-n-butyl phthalide (Zheng *et al*, 1993).



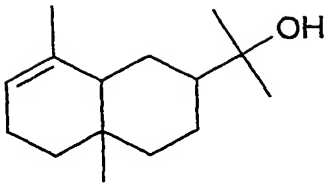

Experimental data

EI-MS: m/z 190.2 (calculated for $C_{12}H_{14}O_2$). $^1$H NMR ($CDC1_3$): $\delta$ 0.85 (t, 3H J=7.1, H-4'), 1.2-2.10 [m, 6H, H-(1' 2', 3')], 5.42 (dd, 1H, J=7.8 and 4.1 Hz, H-3), 7.39 (d, 1H, J=7.5, H-4), 7.46 (t, 1H, J=7.5, H-6), 7.62 (t, 1H, J=7.5 HZ, H-5), 7.83 (d, 1H, J=7.5 Hz, H-7); $^{13}$C NMR ($CDCL_3$): $\delta$14.08 (C-4'), 22.65 (C-3'), 27.01 (C-1'), 34.62 (C-2'), 81.75 (C-3), 121.68 (C-4), 125.57 (C-6), 125.96 (C-9), 128.94 (C-7), 134.20 (C-5), 150.02 (C-8), 171.04 (C-1).

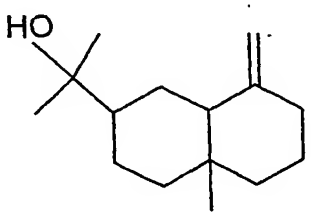
(Large quantity of 6-2 will be obtained by purification using PTLC or semi-preparative HPLC, then $^1$H NMR and $^{13}$C NMR will be acquired again to get clear spectra).

For compound **6-3**, the EI-MS spectrum (Fig.10) showed the molecular ion peak at mass/charge ratio (m/z) 222, corresponding to the molecular formula $C_{15}H_{26}O$. Other major peaks were at m/z 204, 189, 162, 149, 135, 109, 108, 95, 81, 59 and 41. On the basis of EI-MS, the compound 6-3 was identified as mixture of $\alpha$ and $\beta$-Eudesmol (El-Sayed *et al.* 1989).

$^1$H NMR and $^{13}$C NMR spectra will confirm the structure of 6-3. But there is not enough sample by now for measuring $^1$H NMR and $^{13}$ NMR (around 10-20 mg needed). The possible structure of compound 6-3 is as below:



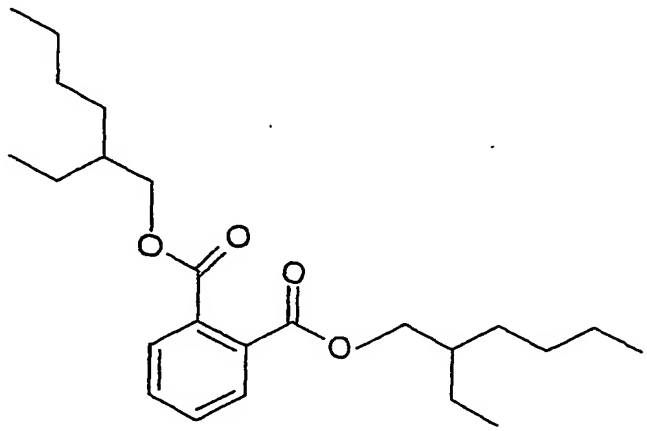

α-Eudesmol




β-Eudesmol

Compound **6-4** was obtained as colourless oil. The EI-MS spectrum of 6-4 (Fig.11) showed the major peaks at m/z 279, 167, 149, 83, 71, 57 and 43. On the basis of EI-MS. the Compound 6-4 was identified as dioctyl phthalate, corresponding to the molecular formula $C_{24}H_{38}O_4$ (MW = 390.54 ) (MS library).

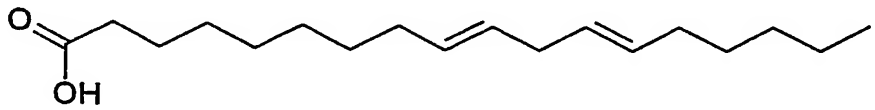
$^1H$ NMR and $^{13}C$ NMR spectra will confirm the structure of 6-4. But there is not enough sample by now for measuring $^1H$ NMR and $^{13}C$ NME (around 10-20 mg needed). The possible structure of compound 6-4 is as below:



Compound **10-1** was obtained as a colourless oil. The EI-MS spectrum (Fig.12) of 10-1 showed the molecular ion peak at mass/charge ration (m/z) 280, corresponding to the molecular formula $C_{18}H_{32}O_2$. Other major peaks were at m/z 137, 123, 109, 95, 81, 67, 55, 54 and 41. On the basis of EI-MS, the compound 10-1 was identified as linoleic acid (cis, cis - 9,12- Octadecadienoic acid) (MS library).

$^1H$ NMR and $^{13}C$ NMR spectra will confirm the structure of 10-1. But there is not enough sample for measuring $^1H$ NMR and $^{13}C$ NMR (around 10-20 mg).

The possible structure of compound 10-1 is as below:




**Conclusion**

Overall the CSE has shown interesting antimicrobial activity against *H. pylori.* Five compounds have been purified which are partly responsible for the antimicrobial properties. The structure elucidation of compounds is still undergoing. Further work will continue to purify the active constituents in subfraction 10 and other subfractions and to test the anti-cytokine activity and cartilage protection properties. If the compounds from subfractions 6 and 10 are not responsible for the anti-inflammatory activity, the constituents maybe reside in other fractions and subfractions.

## References

Bjeldanes L.F. and KIM I.S. (1977) Phthalide components of celery essential oil. J.Org. Chem. 42(13), 23333-5.

Chan. F.K.L., Sung J.Y., Leung V.K.S. *et al* (1997) Randomized trial of eradication of *H. pylori* before non-steroid anti-inflammatory drug therapy to prevent peptic ulcer. Lancet 350, 975-9.

Dey P.M. and Harborne J.B. (1991) Methods in plant Biochemistry, volume 7, Terpenoids, Edited by Charlwood B.V. and Banthorpe D.V., Academic Press, p.65.

El-Sayed A.M., Al-Yahya M.A. Hassan, M.M. (1989) Chemical composition and antimicrobial activity of the essential oil of Chenopodium botrys growing in Saudi Arabia. Int. J. Crude Drug Res. 27, 185-188.

Kitsos C.M. and Stadtländer C.T., (1998) *Helicobacter pylori* in liquid culture: Evaluation of growth rates and ultrastructure. Curr. Microbiol. 37, 88-93.

Lewis D.A., Tharib S.M. and Veitch G.B.A. (1985). The anti-inflammatory activity of celery *Apium graveolens* L. (Fam. Umbelliferae) Int. J. Crude Drug Res. 23, 27-32.

Momin R.A. and Nair M.G. (2001) Mosquitocidal, nematicidal and antifungal compounds from *Apium graveolens* L. seeds. J. Agric. Food Chem. 49, 142-145.

Morgan D., Freedman R., Depew C., and Kraft W. (1987) Growth of *Campylobacter pylori* in liquid media. J. Clin. Microbiol. 25, 2123-2125.

Osato M.S., Reddy S.G. and Graham, D.Y. (1999) Osmotic effect of honey on growth and viability of *H.pylori*. Dig. Dis. Sci. 44, 462-464.

Zheng G.Q.; Kenney P.M.; Zhang J.; Lam L.K.T. (1993) Chemoprevention of benzo[a]pyrene-induced forestomach cancer on mice by natural phthalides from celery seed oil. Nutr. Cancer 19(1), 77-86.

**Table 1.** Effect of the crude extract of CSE on the growth of different strains (3330, 3336 and 3339) on *H. pylori*.

| Strains | MIC (μg/ml) | MBC (μg/ml) |
|---|---|---|
| 3330 | 250 | 500 |
| 3336 | 125 | 500 |
| 3339 | 125 | 500 |

**Table 2** Distribution of antimicrobial activty against *H. pylori* (strain 3339) in the crude extract and different fractions of CSE.

| Fractions | MIC (μg/ml) | MBC (μg/ml) |
|---|---|---|
| Crude extract | 125 | 500 |
| Pet. ether | **15.625** | **31.25** |
| Diethyl ether | 125 | 500 |
| Ethylacetate | >500 | >500 |
| Water | >500 | >500 |

**Table 3** Antimicrobial activity of the subfractions from pet. ether fraction against *H.pylori* (strain 3339).

| Fractions and subfractions | MIC (μg/ml) |
|---|---|
| Pet. ether | **15.625** |
| Sub-1 | >125 |
| Sub-2 | >125 |
| Sub-3 | 125 |
| Sub-4 | 62.5 |
| Sub-5 | 62.5 |
| Sub-6 | **15.625** |
| Sub-7 | 31.25 |
| Sub-8 | 31.25 |
| Sub-9 | 62.5 |
| Sub-10 | **15.625** |
| Sub-11 | 31.25 |

**Table 4** Antimicrobial activities of compounds from subfractions 6 and 10 against *H.Pylori* (strain 3339).

| Compounds | MIC ( g/ml) | MBC ( g/ml) |
|---|---|---|
| sedanenolide | 31.25 | 62.5 |
| 3-n Butyl phthalide | 15.625 | N.T. |
| Eudesmol | 15.625 | 125 |
| Eudesmol + sedanenolide (major) (minor) | 15.625 | N.T. |
| Eudesmol + sedanenolide (minor) (major) | 31.25 | N.T. |
| Linoleic acid | 62.5 | >125 |
| 10-2, 10-3 and 10-4 | 12.5 | 25 |

N.T. : not tested

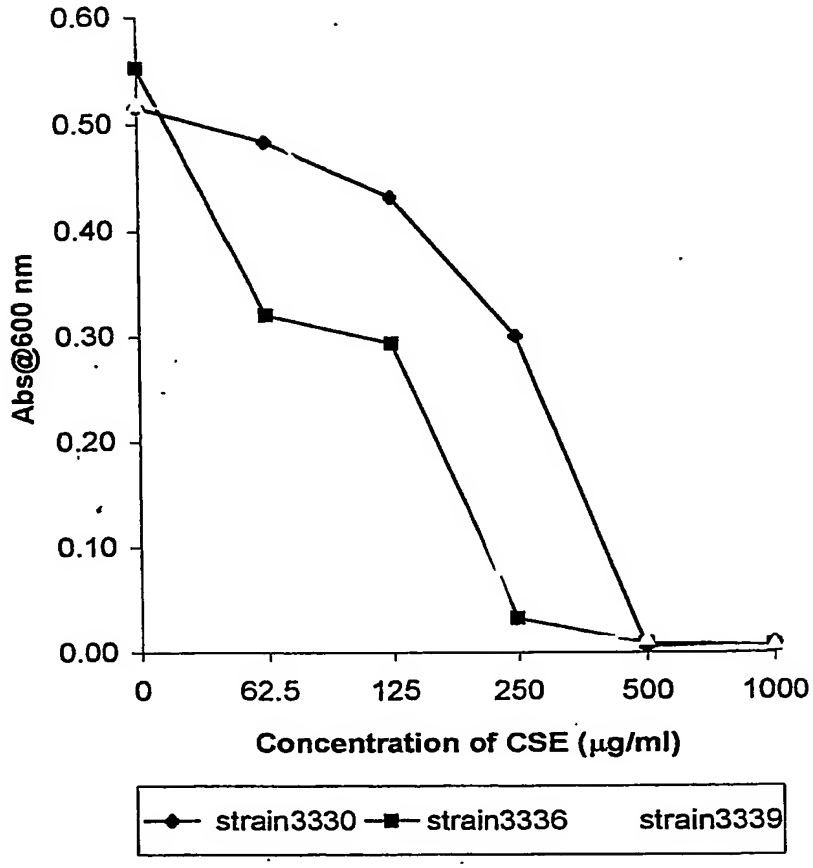



**Fig.1** Effect of CSE crude extract on the growth of the strains (3330, 3336, 3339) of *H.pylori*

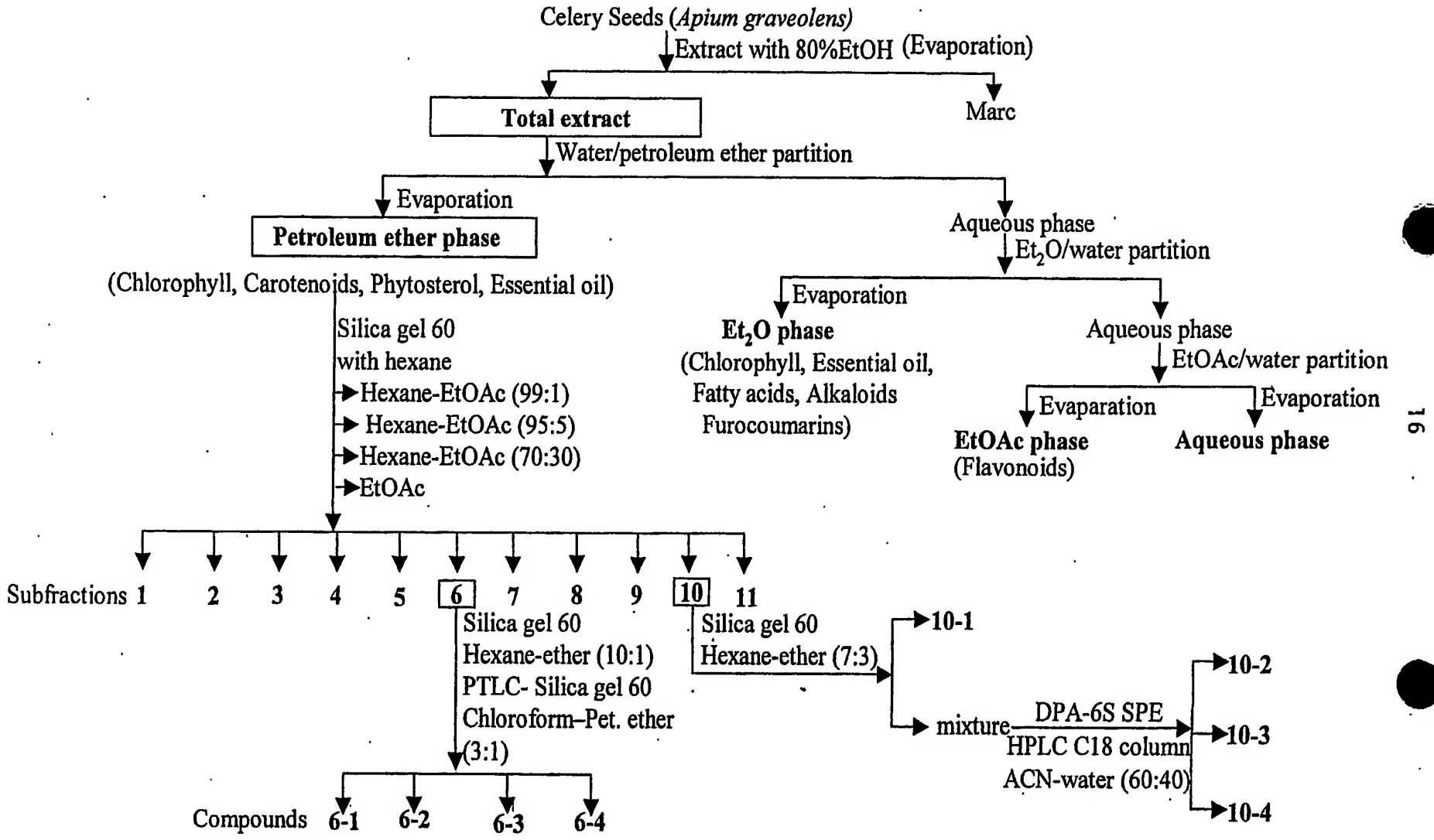



**Fig.2 Bioassay-guided fractionation scheme of celery seed extract (antimicrobial agents enclosed in boxes)**

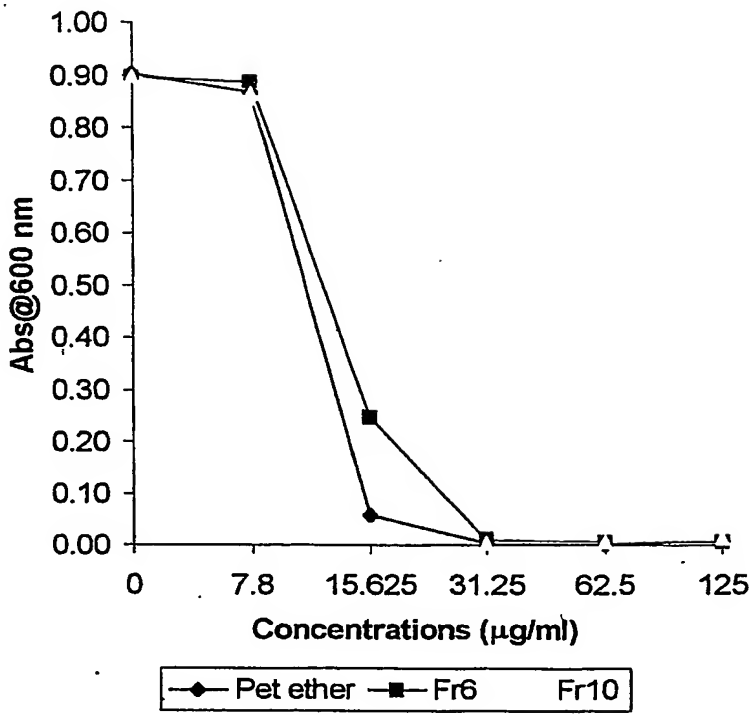



**Fig.3** Antimicrobial activity of pet.ether fraction and subfractions 6 and 10 against *H.pylori* (strain 3339)

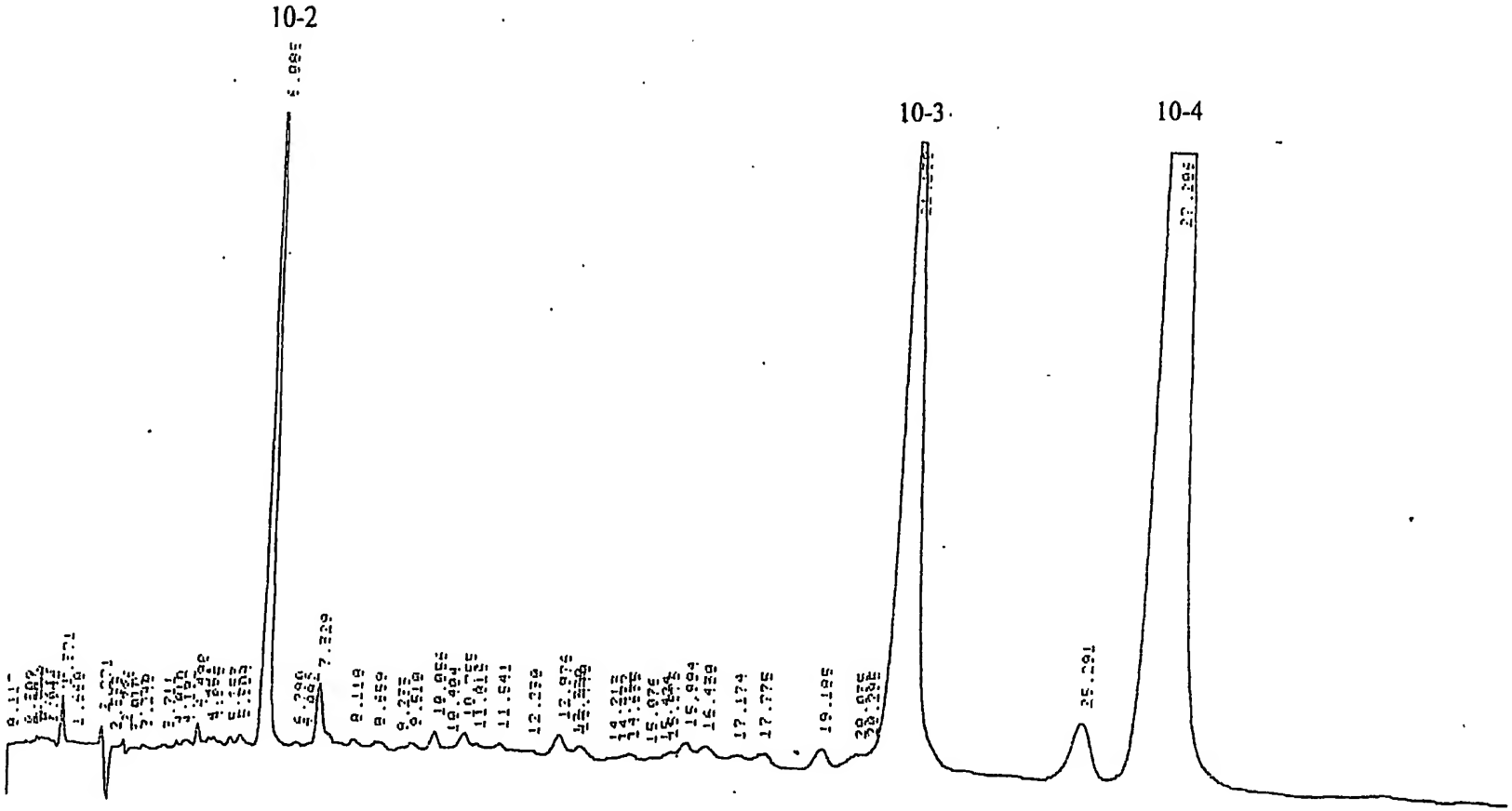



Fig.4 Analytical separation of mixture from subfraction 10. Column: Nucleosil® C18, 250 × 4.6 mm I.D.; Mobile phase: ACN/water (60:40); Flow rate: 1.0 ml/min; Detection: UV @ 236 nm; Injection volume: 10 µl; Sample: 500 µg in 1 ml of 40% ACN in water ; Temperature: Ambient; ATT: 3.

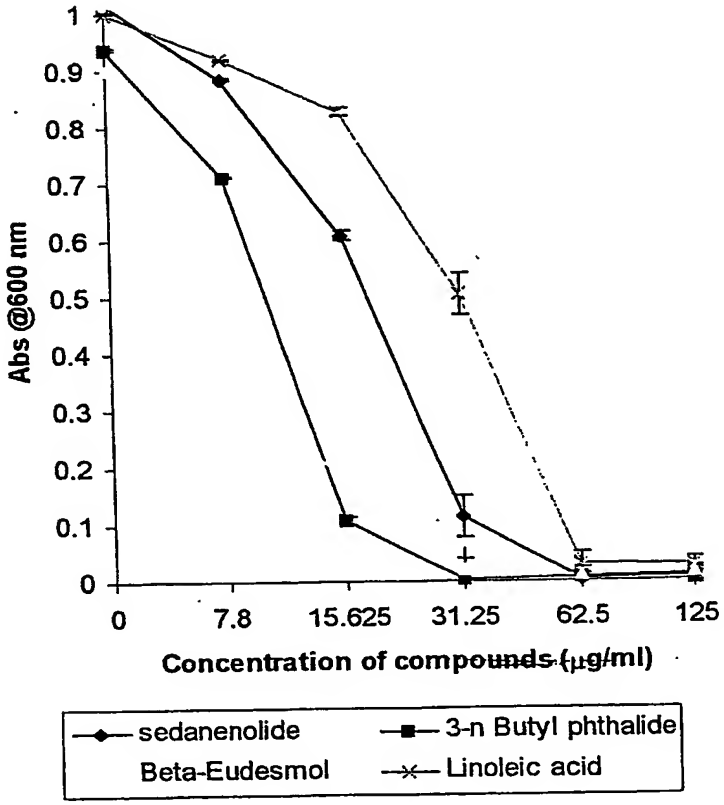



**Fig.5** Antimicrobial activities of compounds against *H. pylori* (strain 3339)

**Fig.6** EI-MS spectrum of compound 6-1

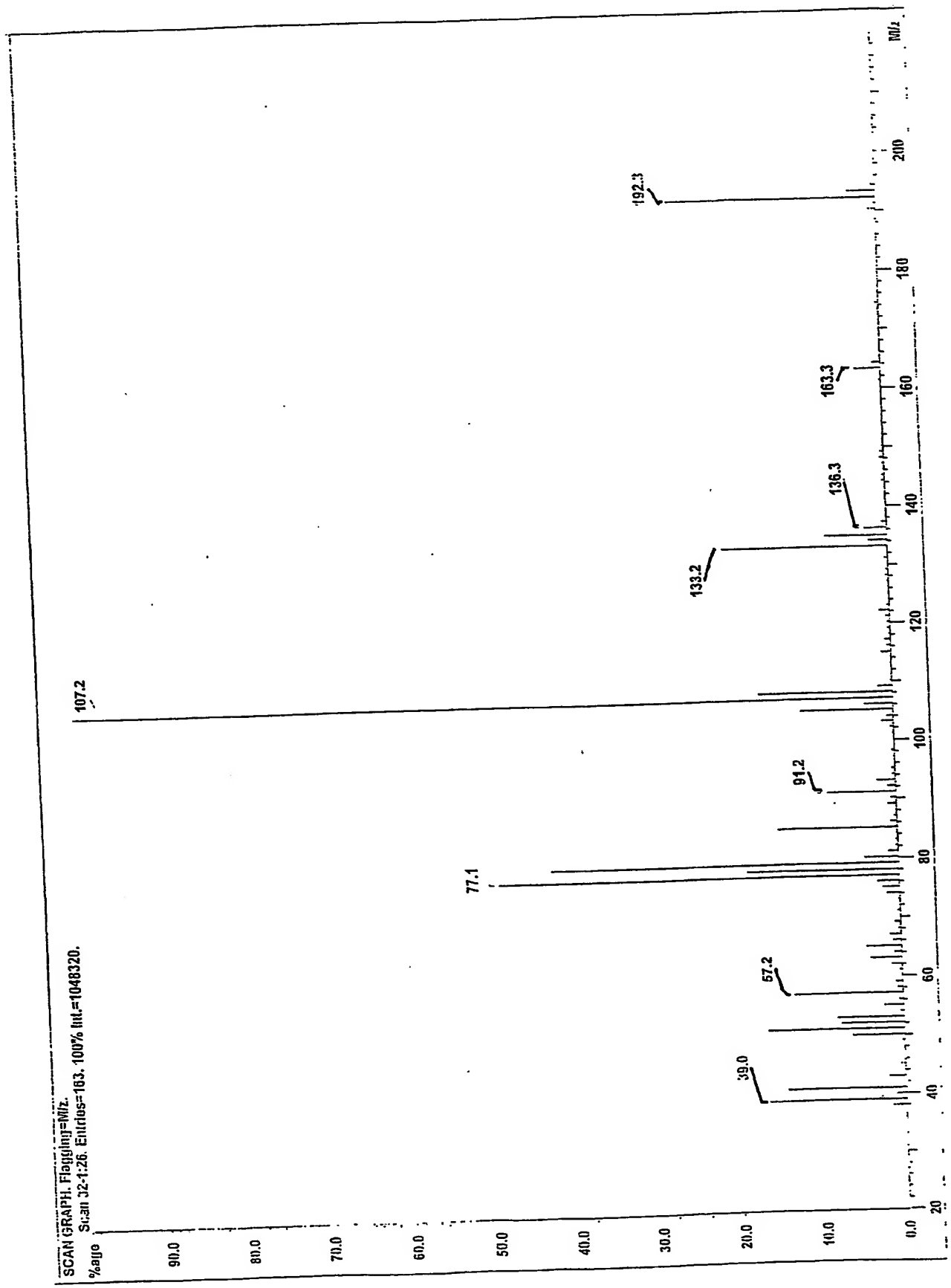


**Fig.7** $^{1}$H NMR spectrum of compound 6-1

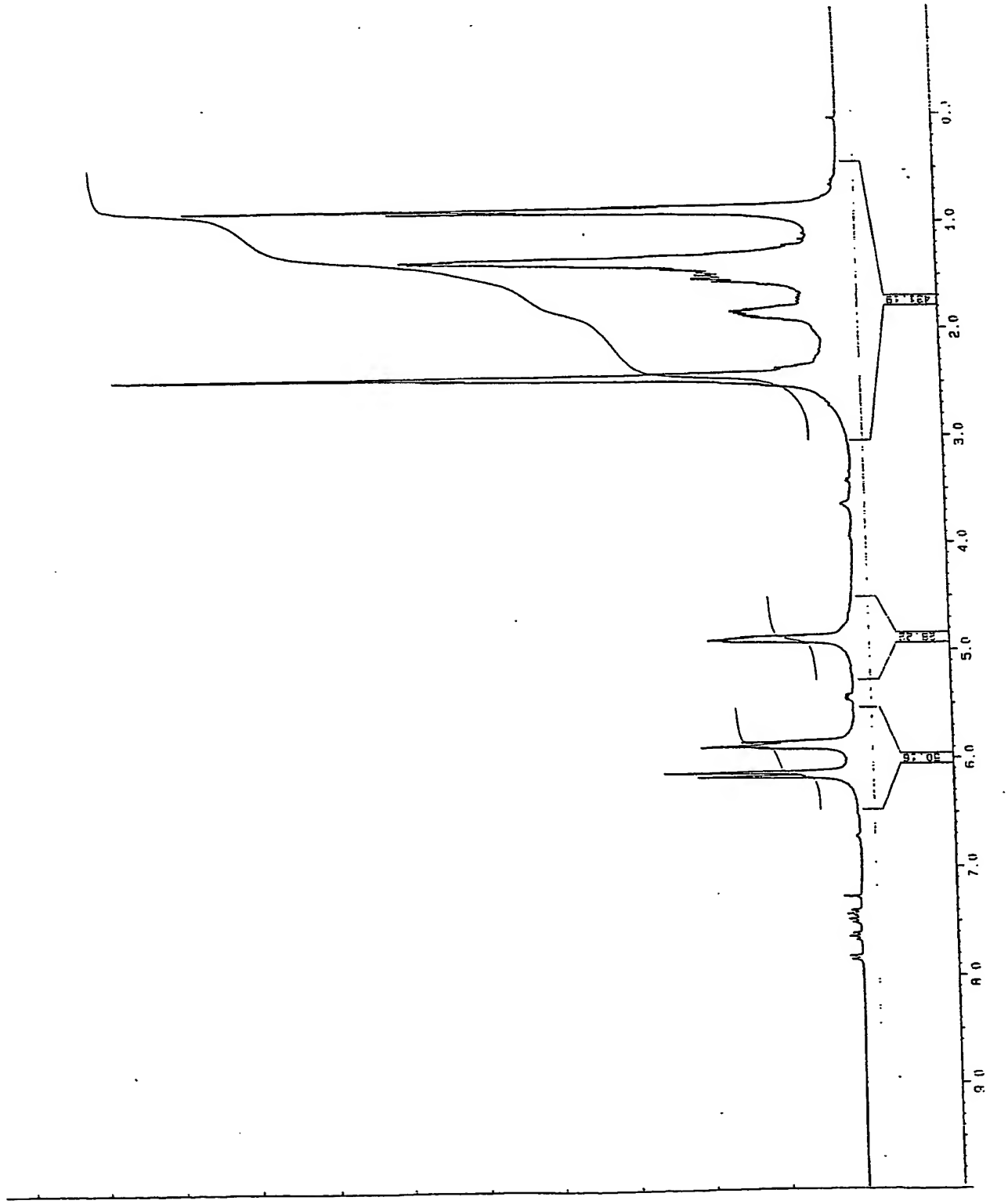


**Fig.8** $^{13}C$ NMR spectrum of compound 6-1

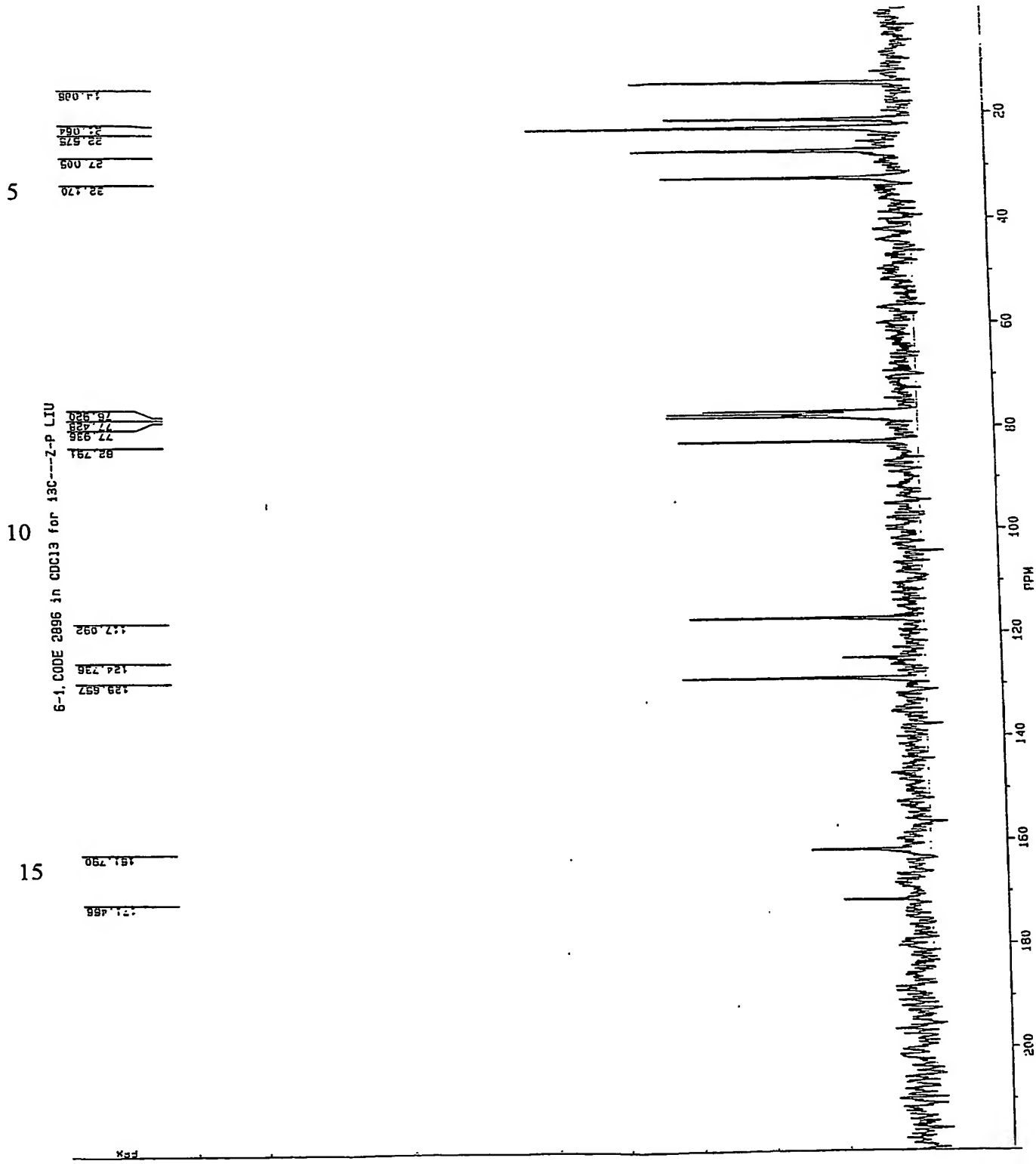


**Fig.9** EI-MS spectrum of compound 6-2

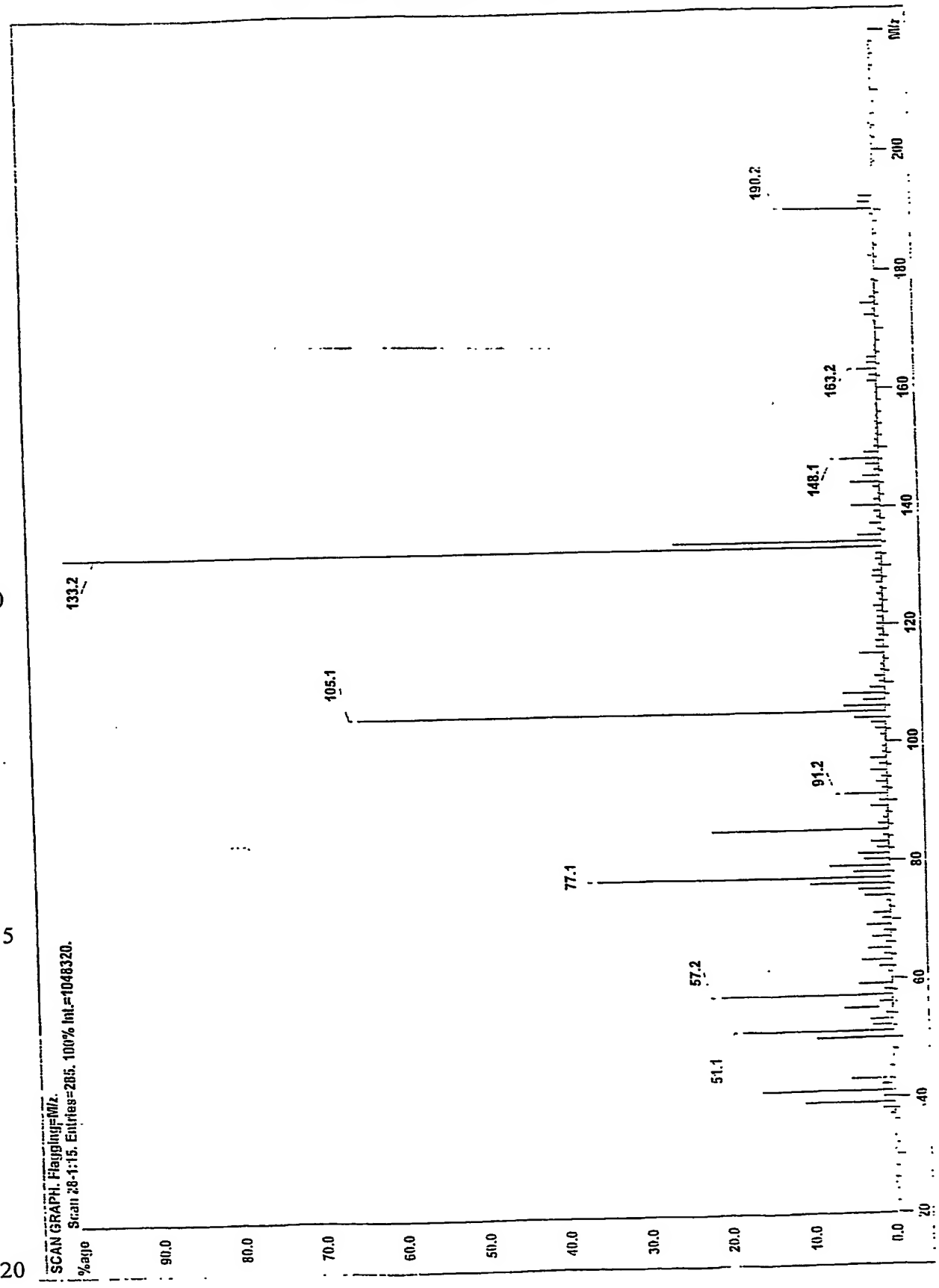


**Fig.10** EI-MS spectrum of compound 6-3

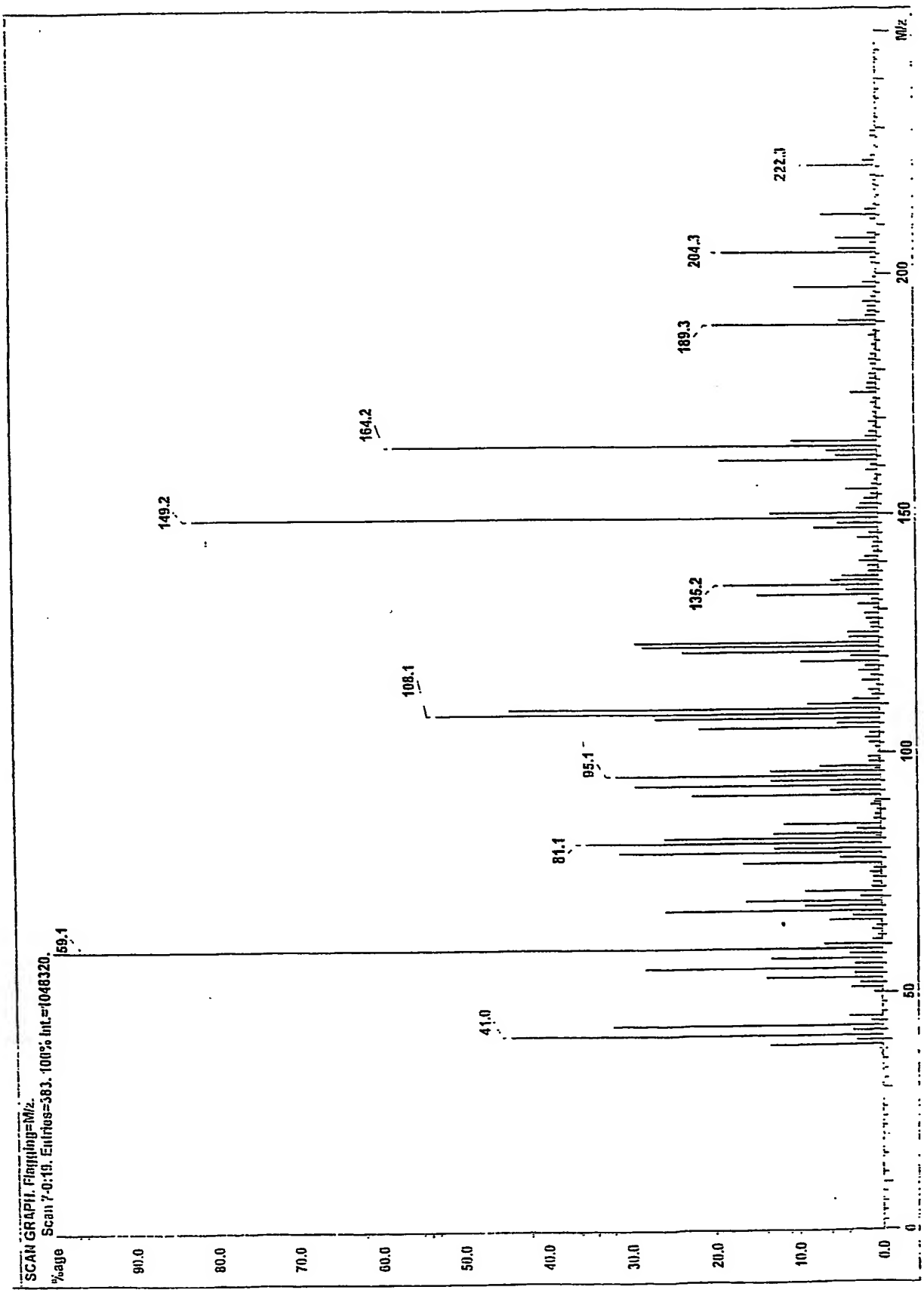


**Fig.11** EI-MS spectrum of compound 6-4

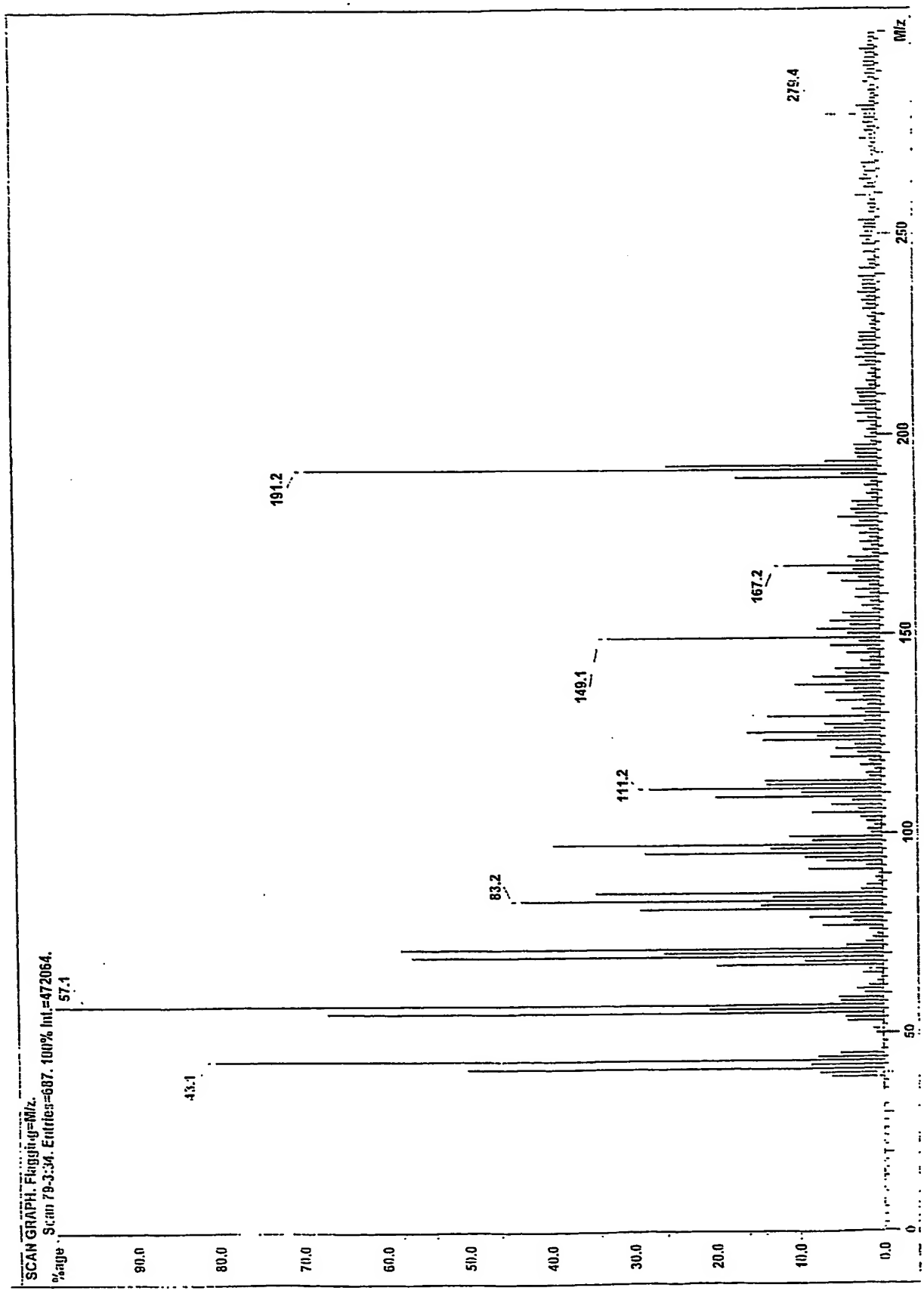
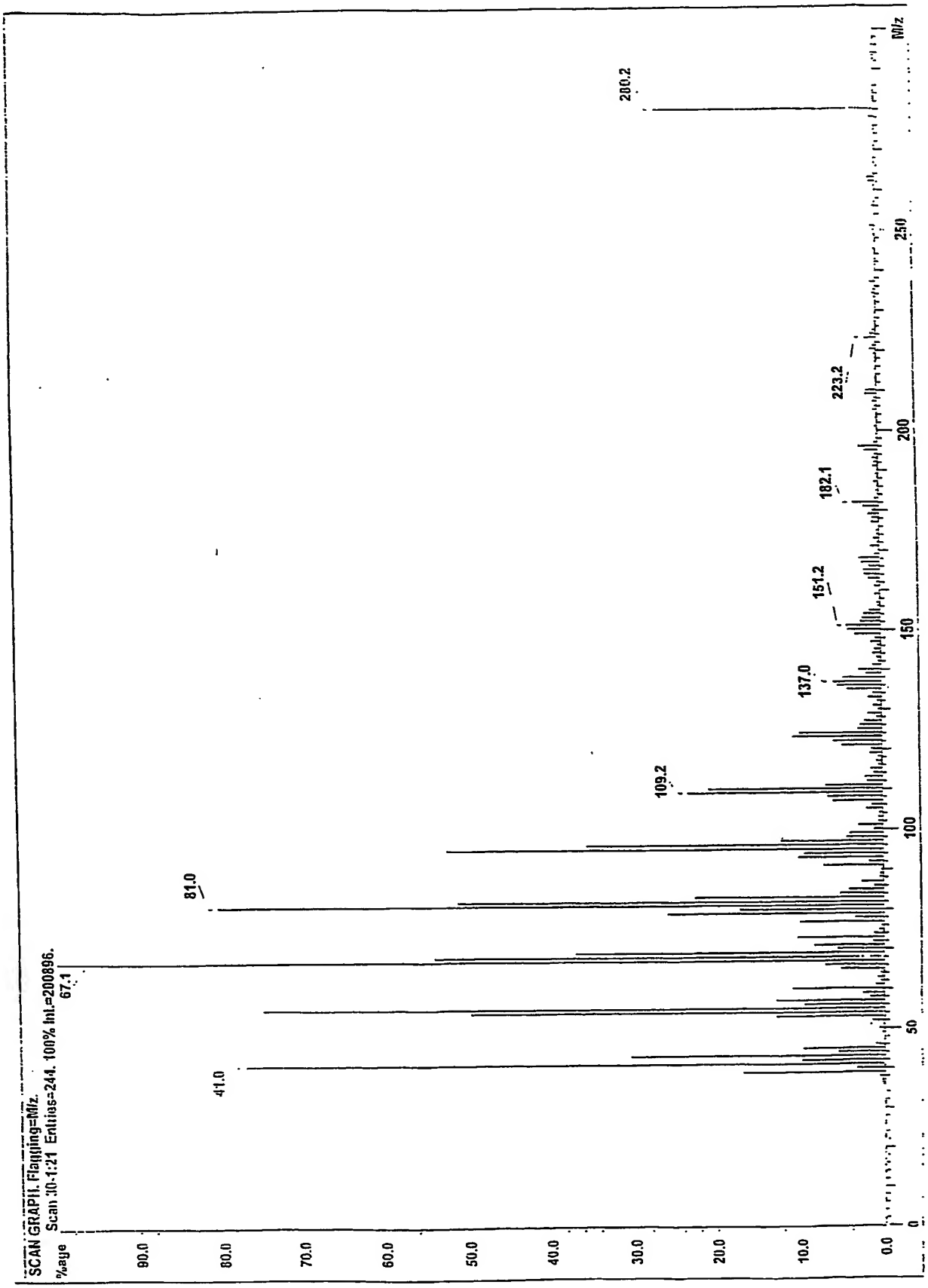


**Fig.12** EI-MS spectrum of compound 10-1



**Patents Form 1/77**

[redacted]nts Act 1977
[redacted]ie 16)

The Patent Office
PATENTS • DESIGNS • COPYRIGHT • TRADE MARKS

[redacted]02 E714662-1 D10100
P01/7700 0.00-0209723.6

THE PATENT OFFICE
A
27 APR 2002
NEWPORT

# Request for grant of a patent

*(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)*

**The Patent Office**

Cardiff Road
Newport
South Wales
NP10 8QQ

1. Your reference — P/184/GBA

2. Patent application number *(The Patent Office will fill in this part)* — 0209723.6 — 27 APR 2002

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*

   8372922001

   Patents ADP number *(if you know it)*

   If the applicant is a corporate body, give the country/state of its incorporation

   The Secretary of State for Environment, Food & Rural Affairs
   (DEFRA), Nobel House, 17 Smith Square, London SW1P 3JR, GB
   acting through the
   Veterinary Laboratories Agency of New Haw, Addleston, Surrey, KT15 3NB
   GB

4. Title of the invention

   Diagnostic Kit

5. Name of your agent *(if you have one)*

   "Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

   Carol P. Greaves et al.

   Greaves Brewster
   24A Woodborough Road
   Winscombe
   North Somerset
   BS25 1AD
   GB

   Patents ADP number *(if you know it)* 7885908001

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number

| Country | Priority application number *(if you know it)* | Date of filing *(day / month / year)* |
|---|---|---|
| | | |

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

| Number of earlier application | Date of filing *(day / month / year)* |
|---|---|
| | |

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:*
   *a) any applicant named in part 3 is not an inventor, or*
   *b) there is an inventor who is not named as an applicant, or*
   *c) any named applicant is a corporate body.*
   *See note (d))*

   Yes

**Patents Form 1/77**

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

| | |
|---|---|
| Continuation sheets of this form | |
| Description | 31 |
| Claim *(s)* | 6 |
| Abstract | 1 |
| Drawing *(s)* | 184 18 |

CF

10. If you are also filing any of the following, state how many against each item.

| | |
|---|---|
| Priority documents | |
| Translations of priority documents | |
| Statement of inventorship and right to grant of a patent *(Patents Form 7/77)* | |
| Request for preliminary examination and search *(Patents Form 9/77)* | |
| Request for substantive examination *(Patents Form 10/77)* | |
| Any other documents *(please specify)* | |

11. I/We request the grant of a patent on the basis of this application.

Signature [signature] Date 26/04/02

12. Name and daytime telephone number of person to contact in the United Kingdom: Carol Greaves 01934 844419

**Warning**

*After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.*

**Notes**

a) *If you need help to fill in this form or you have any questions, please contact the Patent Office on 08459 500505.*

b) *Write your answers in capital letters using black ink or you may type them.*

c) *If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.*

d) *If you have answered 'Yes' Patents Form 7/77 will need to be filed.*

e) *Once you have filled in the form you must remember to sign and date it.*

f) *For details of the fee and ways to pay please contact the Patent Office.*

# DIAGNOSTIC KIT

The present invention relates to antigens derived from the RD1 or RD2 regions of the *Mycobacterium tuberculosis, Mycobacterium bovis* or *Mycobacterium africanum* genomes. Such antigens may be used as diagnostic reagents.

In particular, the present invention relates to diagnostic kits comprising such antigens for differentiating between those mammals infected by tuberculosis, those which have been vaccinated against tuberculosis, and those mammals, which have been sensitised by environmental *Mycobacteria.*

The present invention further relates to novel *Mycobacterium tuberculosis, Mycobacterium bovis* and *Mycobacterium africanum* peptides derived from such antigens.

The present invention also relates to vaccines against *Mycobacterium* infections, in particular, *Mycobacterium tuberculosis, Mycobacterium bovis* or *Mycobacterium africanum* infections, as well as to veterinary and pharmaceutical compositions containing these and their preparations.

Bovine tuberculosis (BTB) is caused by *Mycobacterium bovis* and shares greater than 99.9% DNA identity with *M. tuberculosis,* the main cause of human tuberculosis. Moreover, BTB is a zoonotic disease and was responsible during the 1930s and 1940s for approximately 6% of the total human deaths due to TB, and more than 50% of all cervical lymphadentitis cases in children [Hardie, 1992 #54]. The introduction of pasteurisation of milk in the 1930s dramatically reduced the transmission from cattle to man [Hardie, 1992 #54]. However, it still remains a small but significant cause of human morbidity and mortality especially in developing countries. [Cosivi, 1995 #23] and is seen as one of the most important infectious diseases of both man and other animals in the world.

*Mycobacterium bovis* causes disease in both cattle and man. In the absence of a BTB control programme, TB in cattle can have severe implications for animal welfare, causing reduced productivity and premature death, resulting in substantial economic

losses to affected farms. A compulsory eradication programme based upon the slaughter of infected animals, detected by the single intradermal comparative tuberculin skin test, began in Great Britain (GB) in 1950 and by 1960 it had been implemented in all of Britain. These measures resulted in the dramatic reduction of bovine tuberculosis in GB from incidence rates of around 40% of cattle infected with *M. bovis* to 0.41% in 1996 [Krebs, 1997 #67]. However, despite continued implementation of these control measures, the incidence of BTB in cattle has been steadily rising since 1988, possibly due to a wildlife reservoir of *M. bovis* [Krebs, 1997 #67].

BCG is an attenuated strain of *M. bovis*, and is currently the only available vaccine for the prevention of BTB. Encouraging results with BCG have been reported in New Zealand where a significant level of protection in BCG vaccinated cattle against experimental *M. bovis* infection has been recently demonstrated [Buddle, 1995 #16; Buddle, 1995 #17].

Immunity to *M. tuberculosis* is characterised by three basic features: 1) living bacilli which efficiently induce a protective immune response; 2) specifically sensitised T lymphocytes which mediate this protection, and 3) interferon gamma (IFN-γ) which is an important mediator molecule.

Cattle with a mycobacterial infection will predominantly mount a cellular immune response [Buddle, 1995 #17]. Therefore, the skin test using tuberculin PPD has become an integral part of the bovine tuberculosis eradication programme. In addition to skin tests, blood-based diagnostic assays that measure antigen-induced lymphokine production such as the IFN-γ are also under consideration [Wood, 1994 #136]. The cytokine IFN-γ appears to be critical in the development of immunity to *M. tuberculosis*. For example, both mice with a disruptive IFN-γ gene and humans with mutated IFN-γ receptor are highly susceptible to mycobacterial infections [Flynn, 1993 #188; Cooper, 1993 #186; Jouanguy, 1996 #62]. However, specificity constraints are associated with the use of PPD in such assays. These arise due to the crude mixture of *M. bovis* proteins that it contains, many that are cross-reactive with other environmental mycobacterial species, e.g., *M. avium or M. intracellulare* and

importantly the vaccine strain *M. bovis* Bacille Calmette-Guerin (BCG) [Buddle, 1995 #16; Berggren, 1981 #9; Hubrig, 1958 #59].

A cattle vaccine would reduce the risk of cattle infection and hence result in lower tuberculin test frequencies and significant cost savings. It is believed that the development of an improved cattle vaccine holds the best long-term prospect for BTB control in British herds [Krebs, 1997 #67]. In addition, it would be desirable to develop a complimentary diagnostic test to differentiate between vaccinated animals and those infected with M. bovis (*differential diagnosis*) in parallel with the vaccine to ensure continuation of the test and slaughter-based control strategies [Krebs, 1997 #67].

As previous studies have demonstrated, diagnostic reagents which distinguish between vaccinated and infected cattle can be developed using specific, defined antigens that are present in virulent *M. bovis* but absent from the vaccine strain [Buddle, 1999 #18; Vordermeier, 2000 #126; Vordermeier, 2001 #169]. Genetic analysis of BCG has revealed that several large genomic regions have been deleted during attenuation and subsequent prolonged propagation in culture [Behr, 1999 #8; Gordon, 1999 #43]. These regions have been characterised and antigens from one of these regions, RD1 [Mahairas, 1996 #83], have been studied extensively in several species including humans and cattle [Lalvani, 2001 #174; Pollock, 1997 #102]. For example, it has been recently demonstrated that protein or peptide cocktails composed of two RD1 region antigens, ESAT-6 and CFP-10, can be used to distinguish between BCG vaccinated and M. bovis infected cattle [van pinxteren, 2000 #119; Vordermeier, 2001 #169]. However, the level of sensitivity achieved with these antigens has not reached that of tuberculin. It would, therefore, be desirable to provide other antigens in order to achieve this desired sensitivity. Such antigens may also be useful in supplementing the ESAT-6 and CFP-10 to achieve even greater sensitivity.

In alternative approach to using recombinant proteins is the application of overlapping synthetic peptides derived from those antigens described above. Synthetic peptides have the advantages of lower production costs, easier standardisation, improved quality control and carry no risk of infection since they are chemically synthesised.

Such synthetic peptide epitopes have been found to have great potential in the study of immune responses in cattle and in the development of diagnostic reagents. For example, formulation of 10 synthetic peptides derived from ESAT-6 and CFP-10 resulted in similar cellular immune responses to those seen with the whole recombinant antigens. When assayed in cattle this cocktail could distinguish between *M. bovis* infected animals and BCG vaccinated cattle with sensitivity similar to PPD and with a greater specificity [Vordermeier, 2001 #169].

Differential diagnosis is not the only concern associated with BCG. BCG vaccination studies have highlighted the variability with regard to its efficacy. In humans, this ranges from 0 to 80% when tested in different populations, with consistently poor results observed in the equatorial regions [Fine, 1989 #37]. Similar variations in efficacy have also been reported in BCG vaccination experiments and trials in cattle (e.g. [Buddle, 2002 #184; Berggren, 1981 #9; Buddle, 1995 #16; Buddle, 1995 #17; Hubrig, 1958 #59]). It would therefore be desirable to improve or supplement BCG vaccination. Strategies to generate novel tuberculosis vaccines include sub-unit vaccination with either DNA vaccines or protein subunits (Rev. [Andersen, 2001 #173]). Antigens such as MPT-64 and ESAT-6 (Rev. [Anderson, 2001 #173]), whose genes were deleted in BCG, have been tested as DNA vaccines and imparted protective immunity in small animal models.

The present invention seeks to provide an improved diagnostic test to differentiate between vaccinated mammals and those infected with tuberculosis. Preferably, the test of the present invention differentiates between animals vaccinated against *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobacterium africanum* and those infected with *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobacterium africanum*.

The present invention also seeks to provide an improved vaccine for control of tuberculosis and in particular to control tuberculosis in cattle. The tuberculosis disease also affects a number of other different animal species such as guinea pigs, badgers, possums and deer. The vaccines of the present invention may therefore be useful in the control of tuberculosis infections in such different animals.

According to a first aspect of the present invention there is provided a diagnostic agent comprising a polypeptide derived from the RD1 or RD2 regions of the *Mycobacterium tuberculosis, Mycobacterium bovis* or *Mycobacterium africanum* genomes, or a variant, homologue or fragment of these, with the proviso that the polypeptide is not a ESAT-6, CFP-10, MPT-64 or a polypeptide encoded by the Rv1984c, Rv3871, Rv3872 or Rv3873 regions of the *Mycobacterium tuberculosis, Mycobacterium bovis* or *Mycobacterium africanum* genomes.

The term "polypeptide" as used herein includes long chain peptides including proteins and epitopic fragments thereof. Such polypeptides generally comprise one or more chains of amino acids joined covalently through peptide bonds and are typically greater than 10,000 MW. Also included are oligopeptides comprising three or more amino acid residues covalently linked through peptide bonds.

The polypeptide is preferably derived from the *Mycobacterium tuberculosis* genome and a member of the PE/PPE protein family.

The term "derived from" as used herein means any polypeptide or peptide encoded by an open reading frame from the RD1 or RD2 regions of the *Mycobacterium tuberculosis, Mycobacterium bovis* or *Mycobacterium africanum* genomes. Also included are fragments of the open reading frames and variants thereof as long as such fragments are still capable of being used as diagnostic reagents.

The polypeptide preferably comprises the sequence shown in SEQ ID Nos 1, 3,4 and 6, or a fragment, homologue or variant thereof.

The term "fragment thereof" as used herein in relation to an amino acid sequence refers to any portion of the given amino acid sequence which has the same activity as the complete amino acid sequence. Fragments will suitably comprise at least 10 and preferably at least 20 consecutive amino acids from the basic sequence. Preferably, the sequence comprises 17 amino acids. Fragments of the polypeptide include deletion mutants and polypeptides where small regions of the polypeptides are joined together. The fragments preferably contain at least one antigenic region.

The term "variant thereof" as used herein in relation to an amino acid sequence means sequences of amino acids which differ from the base sequence from which they are derived in that one or more amino acids within the sequence are substituted for other amino acids. Amino acid substitutions may be regarded as "conservative" where an amino acid is replaced with a different amino acid with broadly similar properties. Non-conservative substitutions are where amino acids are replaced with amino acids of a different type.

By "conservative substitution" is meant the substitution of an amino acid by another one of the same class; the classes being as follows:

| CLASS | EXAMPLES OF AMINO ACID |
|---|---|
| Nonpolar: | Ala, Val, Leu, Ile, Pro, Met, Phe, Trp |
| Uncharged polar: | Gly, Ser, Thr, Cys, Tyr, Asn, Gln |
| Acidic: | Asp, Glu |
| Basic: | Lys, Arg, His |

As is well known to those skilled in the art, altering the primary structure of a peptide by a conservative substitution may not significantly alter the activity of that peptide because the side-chain of the amino acid which is inserted into the sequence may be able to form similar bonds and contacts as the side chain of the amino acid which has been substituted out. This is so even when the substitution is in a region, which is critical in determining the peptides conformation.

Non-conservative substitutions are possible provided that these do not interrupt with the antigenicity of the polypeptide.

Broadly speaking, fewer non-conservative substitutions will be possible without altering the biological activity of the polypeptide. Suitably, variants will be at least 50% identical, 60% identical, preferably at least 75% identical, and more preferably at least 90% identical to the base sequence.

Homology in this instance can be judged for example using the algorithm of Lipman-Pearson, with Ktuple:2, gap penalty:4, Gap Length Penalty:12, standard PAM scoring matrix (Lipman, D.J. and Pearson, W.R., Rapid and Sensitive Protein Similarity Searches, *Science*, 1985, vol. 227, 1435-1441).

Where a differential diagnostic test is to be carried out, the diagnostic reagent preferably comprises the sequence shown in SEQ ID Nos 1 or 3, or a fragment, homologue or variant thereof. These polypeptides are used to differentiate between tuberculosis-infected and tuberculosis vaccinated mammals.

The diagnostic reagent used in the differential diagnostic test preferably differentiates between *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobacterium africanum* -infected mammals and mammals vaccinated against *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobacterium africanum.*

Alternatively, the diagnostic reagent according to the present invention comprises SEQ ID Nos 3 or 6, or a fragment, homologue or variant thereof. In this instance, the polypeptides are used as diagnostic reagents, which differentiate between mammals, which are either vaccinated against or infected by tuberculosis and mammals, sensitised by environmental mycobacteria.

The diagnostic reagent used in a specific diagnostic test preferably differentiates between *Mycobacterium bovis* -infected and mammals sensitised by environmental mycobacteria.

According to a second aspect of the present invention there is provided a peptide derived from an RD1 or RD2 region of the *Mycobacterium tuberculosis, Mycobacterium bovis* or *Mycobacterium africanum* genomes, or a variant, homologue or fragment of these, with the proviso that the peptide is not derived from a ESAT-6 or CFP-10 polypeptide.

Such peptides are also capable of being used as diagnostic reagents and are preferably synthetic peptides having the advantages discussed above. One such peptide is a

peptide derived from SEQ ID NO.5, which is shown in Figure 6 as SEQ ID NO 7. Fragments, homologues and variants of this peptide are also included herein.

The term "peptide" as used herein includes small proteins (generally less than about 10,000 MW). The peptides of the present invention generally comprise two or more amino acid residues linked together covalently through peptide bonds.

Such peptides may be used as diagnostic reagents, either on their own or preferably with one or more other peptides according to the present invention in order to achieve greater sensitivity and specificity of a diagnostic test. For example, protein or peptide cocktails composed of antigens from the RD1 or RD2 regions of the *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobacterium africanum* genomes may be used in addition to the antigens of the present invention in order to enhance the specificity of the diagnostic test. In particular, peptide cocktails from the antigens, ESAT-6 and CFP-10 may be used.

The peptides of the present invention may be used in either specific or differential diagnostic tests. The peptide as shown in SEQ ID NO 7 may be used in a specific diagnostic test to differentiate between those mammals, which are either vaccinated against or infected by tuberculosis, and those mammals which have been sensitised by environmental mycobacteria. In particular, the peptide is especially useful in differentiating between *Mycobacterium bovis*-infected mammals, such as cattle or calves, and those animals sensitised by environmental bacteria.

According to a third aspect of the present invention, there is provided a nucleic acid or polynucleotide encoding any one of the polypeptides or peptides of the invention, or a fragment, homologue or variant thereof. The nucleic acid may be DNA or RNA, and where it is a DNA molecule, it may comprise a cDNA or genomic DNA. These nucleic acids may themselves be useful as vaccines and such vaccines form a further aspect of the present invention. Preferably, the nucleic acid comprises the sequence shown in SEQ ID Nos 8, 10, 11 or 13, or a variant or fragment thereof.

The term "fragment thereof" as used herein in relation to a nucleic acid or polynucleotide sequence refers to any portion of the given polynucleotide sequence

which exhibits the same activity as the complete polynucleotide sequence. Fragments will suitably comprise at least 15, preferably at least 30 and more preferably at least 60 consecutive bases from the basic sequence.

The term "variant thereof" in relation to a polynucleotide or nucleic acid sequences means any substitution of, variation of, modification of, replacement of deletion of, or the addition of one or more nucleic acid(s) from or to a polynucleotide sequence providing the resultant protein sequence encoded by the polynucleotide exhibits the same properties as the protein encoded by the basic sequence. The term therefore includes allelic variants and also includes a polynucleotide which substantially hybridises to the polynucleotide sequence of the present invention. Preferably, such hybridisation occurs at, or between low and high stringency conditions. In general terms, low stringency conditions can be defined as 3 x SSC at about ambient temperature to about 55°C and high stringency condition as 0.1 x SSC at about 65°C. SSC is the name of the buffer of 0.15M NaCl. 0.015M tri-sodium citrate. 3 x SSC is three times as strong as SSC and so on.

Typically, variants have 62% or more of the nucleotides in common with the polynucleotide sequence of the present invention, more typically 65%, preferably 70%, even more preferably 80% or 85% and, especially preferred are 90%, 95%, 98% or 99% or more identity.

When comparing nucleic acid sequences for the purposes of determining the degree of identity, programs such as BESTFIT and GAP (both from Wisconsin Genetics Computer Group (GCG) software package). BESTFIT, for example, compares two sequences and produces an optimal alignment of the most similar segments. GAP enables sequences to be aligned along their whole length and fins the optimal alignment by inserting spaces in either sequence as appropriate. Suitably, in the context of the present invention when discussing identity of nucleic acid sequences, the comparison is made by alignment of the sequences along their whole length.

According to a fourth aspect of the present invention there is provided a diagnostic kit comprising at least one polypeptide or peptide encoded by the sequences shown as

SEQ ID Nos 1, 3, 4, 6 and 7, or a fragment, homologue or variant thereof, and optionally at least one polypeptide encoded by the sequences shown by the sequences shown as SEQ ID Nos 2 and 5, and optionally one or more reagents, to differentiate between tuberculosis-infected and tuberculosis-vaccinated mammals. The polypeptide and peptide sequences of the invention suitably provide a means for detecting the recognition of the polypeptides or peptides by the T-cell. Preferably, the diagnostic kit differentiates between *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobacterium africanum* -infected mammals and mammals vaccinated against *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobacterium africanum.*

Where the kit is intended to be used to differentiate between those mammals infected by *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobactrium africanum* and those mammals which have been vaccinated against *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobacterium africanum,* the kit will preferably comprise the polypeptides encoded by the sequences shown as SEQ ID Nos 1, 2 and 3, or a fragment, homologue or variant thereof.

Where the kit is intended to be used to differentiate between those mammals infected by *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobactrium africanum* and mammals sensitised by environmental mycobacteria, the kit will preferably comprise polypeptides or peptides encoded by the sequences shown as SEQ ID Nos 4, 5, 6, and optionally, SEQ ID NO. 7 or a fragment, homologue or variant thereof.

The diagnostic kit may also comprise one or more polypeptides or peptides from the RD1 region of the *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobacterium africanum* genomes. Protein or peptide cocktails composed of such polypeptides may also be used. Especially preferred are peptide cocktails composed of the ESAT-6 and/ or the CFP-10 polypeptides. Such peptide cocktails may be used to enhance the sensitivity of the diagnostic tests of the present invention.

According to a fifth aspect of the present invention, there is provided a method of diagnosing infection in a host, or exposure of a host, to a mycobacterium comprising i) contacting a population of cells from the host with a polypeptide derived from an

RD1 or RD2 region of the *Mycobacterium tuberculosis, Mycobacterium bovis* or *Mycobacterium africanum* genomes, or a variant, homologue or fragment of these, which polypeptide may be used as a diagnostic reagent, with the proviso that the polypeptide is not a ESAT-6, CFP-10, MPT-64 or a polypeptide encoded by the Rv1984c, Rv3871, Rv3872 or Rv3873 regions of the *Mycobacterium tuberculosis, Mycobacterium bovis* or *Mycobacterium africanum* genomes; and ii) determining whether the cells of said cell population recognise the polypeptide or fragment or variant thereof.

The population of cells used in the method is suitably a population of T-cells. The method preferably diagnoses infection by *Mycobacterium tuberculosis, Mycobacterium bovis* or *Mycobacterium africanum.*

According to a sixth aspect of the present invention, there is provided a pharmaceutical or veterinary composition comprising a polypeptide or peptide according to the present invention, or a polynucleotide or nucleic acid encoding such polypeptides or peptides, in combination with a pharmaceutically or veterinarily acceptable carrier.

The carriers may be solid or liquid as understood in the art. They may be obtained by conventional procedures using conventional pharmaceutical excipients, well known in the art.

In particular, the compositions of the invention may be in a form suitable for oral use (for example as tablets, lozenges, hard or soft capsules, aqueous or oily suspensions, emulsions, dispersible powders or granules, syrups or elixirs), for administration by inhalation (for example as a finely divided powder or a liquid aerosol), for administration by insufflation (for example as a finely divided powder) or for parenteral administration (for example as a sterile aqueous or oily solution for intravenous, subcutaneous, intramuscular or intramuscular dosing or as a suppository for rectal dosin.

The pharmaceutical or veterinary compositions are preferably in the form of a sterile injectable aqueous or oily suspension, which may be formulated according to known

procedures using one or more of the appropriate dispersing or wetting agents and suspending agents, which have been mentioned above. A sterile injectable preparation may also be a sterile injectable solution or suspension in a non-toxic parenterally-acceptable diluent or solvent, for example a solution in 1,3-butanediol.

Where the compositions of the invention comprise a nucleic acid, they are preferably formulated for parenteral administration and in particular intramuscular injection, although other means of application are possible as described in the pharmaceutical literature, for example administration using a Gene Gun, (Bennett et al., (2000), Vaccine 18, 1893-1901). Oral or intra-nasally delivered formulations are also possible. Such formulations include delivery of the plasmid DNA via a bacterial vector such as species of Salmonella or Listeria (Sizemore et al (1997). Vaccine 15, 804-807).

Formulation techniques generally are well known and are described for example in Chapter 25.2 in Volume 5 of Comprehensive Medicinal Chemistry (Corwin Hansch; Chairman of Editorial Board), Pergamon Press 1990.

The amount of active ingredient that is combined with one or more excipients to produce a single dosage form will necessarily vary depending upon the host treated and the particular route of administration. Dosage unit forms will generally contain about 1 mg to about 500 mg of an active ingredient.

The size of the dose for therapeutic or prophylactic purposes of the composition of the invention will naturally vary according to the age and sex of the animal or patient and the nature of the active component and the route of administration, according to well known principles of medicine. Generally speaking however, for administration to a human as a prophylactic vaccine, dosage units of from 0.25 μg to 2.5mg will be typically employed.

According to a seventh aspect of the present invention, there is provided a polypeptide or peptide of the present invention for use as a medicament.

The polypeptides or peptides of the present invention are preferably used as vaccines against tuberculosis caused by *Mycobacterium tuberculosis, Mycobacterium bovis* or *Mycobacterium africanum* in a mammal.

According to an eighth aspect of the present invention, there is provided the use of a polypeptide or peptide according to the present invention in the preparation of a vaccine. Most preferably the polypeptide has the sequence shown in SEQ ID NO. 7 or an epitopic fragment thereof.

The polypeptide of the present invention is protective against tuberculosis infection and therefore may be used as a prophylactic or therapeutic vaccine, and these form a further aspect of the invention.

The vaccine is preferably used to vaccinate against tuberculosis. It may be used as a vaccine against tuberculosis in both humans and cattle. It is, however, preferably used as a vaccine in cattle.

Preferably, the vaccine comprises protein subunits. Alternatively, it may comprise subunits of the DNA encoding for the polypeptide.

Alternatively, it may comprise a nucleic acid such as a DNA or cDNA encoding for the subunits. When it comprises a nucleic acid, this is suitably incorporated into an expression vector, in such a way that the protein subunit is expressed in the host animal. For example, the nucleic acid may be incorporated into a virus vector such as a vaccinia or adenovirus vector, or a plasmid to form a so-called "naked DNA" vaccine. The vector may contain the usual expression control functions such as promoters, enhancers and signal sequences, as well as selection markers in order to allow detection of successful transformants. The nature of these will depend upon the preceise nature of the vector chosen and will be known to or readily determinable by a person skilled in the art.

Preferably, vaccine compositions will further comprise an adjuvant such as in order to enhance the immune response to the biologically active material administered. Suitable adjuvants include pharmaceutically acceptable adjuvants such as Freund's

incomplete adjuvant, aluminium compounds and, preferably adjuvants which are known to up-regulate mucosal responses such as CTB, the non-toxic pentameric B subunit of cholera toxin (CT).

According to a nineth aspect of the present invention, there is provided the use of a peptide according to the present invention to produce an antibody specific to the peptide.

According to a tenth aspect of the present invention, there is provided a method of protecting a mammal against infection by *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobacterium africanum* comprising administering to said mammal a polypeptide, peptide or pharmaceutical or veterinary composition according to the present invention which produces an immune response against *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobacterium africanum.*

Polypeptides of the invention may be isolated from strains of *M. bovis, M. tuberculosis* or *M. africanum.* Preferably, they are prepared synthetically using conventional peptide synthesisers. Alternatively, they may be produced using recombinant DNA technology or isolated from natural sources followed by any chemical modification, if required. In these cases, nucleic acids encoding the polypeptides are incorporated into suitable expression vectors, which are then used to transform a suitable host cell, such as a prokaryotic cell such as *E. coli.* The transformed cells are cultured and the polypeptide isolated therefrom. Vectors, cells and methods of this type form further aspects of the present invention.

A highly preferred embodiment of the present invention is a diagnostic kit comprising the polypeptides encoded by the sequences shown as SEQ ID Nos 1 to 3 and further comprising one or more polypeptides derived from the RD1 region of *Mycobacterium tuberculosis, Mycobacterium bovis* or *Mycobacterium africanum,* and optionally one or more reagents, for differentiating between cattle infected by *M. bovis* and cattle which have been vaccinated with BCG or with a vaccine according to the present invention.

A further highly preferred embodiment of the present invention is a diagnostic kit comprising the polypeptides and peptides encoded by the sequences shown as SEQ ID Nos 4 to 7 and further comprising one or more polypeptides derived from the RD1 region of *Mycobacterium tuberculosis, Mycobacterium bovis* or *Mycobacterium africanum,* and optionally one or more reagents, for differentiating between cattle which have either been vaccinated against or infected by *M. bovis* and those cattle which have been sensitised by environmental mycobacteria.

A further preferred embodiment of the present invention is a vaccine comprising a peptide having the sequence shown in SEQ ID No 7.

An advantage of the present invention is that the level of sensitivity achieved in diagnostic tests with these antigens is higher than the sensitivity achieved with the antigens ESAT-6 and CFP-10. In addition, the level of specificity of the antigen of the present invention is higher than that of PPD, which is currently used. PPD has the disadvantage that it cross-reacts with other environmental mycobacterial species and the vaccine strain *M. bovis* Bacille Calmette-Guerin (BCG). Such diagnostic tests will enable the transfer from skin testing regimes to vaccine regimes to be implemented.

A further advantage of the present invention is the provision of a test which can distinguish between those mammals that have been vaccinated against tuberculosis, and in particular *M. bovis,* and those which have been infected with *M. bovis*. This allows the selective slaughter of animals which would appear from current tests to all be infected, thereby saving the lives of many animals.

Cattle models of *M. bovis* infection and BCG vaccination were studied to identify highly immunogenic antigens from three genomic regions deleted in BCG Pasteur (RD1, RD2, RD14) [Behr, 1999 #8, Mahairas, 1996 #83] that could be useful as specific diagnostic reagents or subunit vaccine candidates. Five hundred and thirty six overlapping synthetic peptides derived from the sequence of 13 antigens (open reading frames) encoded in these regions were synthesised and used to diagnose infected or vaccinated cattle. The previously mentioned ESAT-6/CFP-10 peptide

cocktail was also included as a gold standard with which to compare and all tests performed used the BOVIGAM ELISA for the detection of bovine IFN-γ.

The present invention will now be described only by way of example, in which reference shall be made to the Figures, in which:

Figure 1 shows the recognition of RD1 products by a *M. bovis* infected cow (A, C and E) and a PPD-A reactor (B, D and F). Whole blood was cultured in the presence of peptide pools of between 8-11 peptides representing RD1 (A and B), RD2 (C and D) and RD14 (E and F) at 5 each peptide/ml. Dashed horizontal lines indicate positive cut-off ($OD_{450}$ values with antigens minus $OD_{450}$ without antigens $\geq$ 0.1);

Figure 2 shows IFN-γ responses induced by RD region antigens by *M. bovis* infected (22), BCG vaccinated (6) and PP-A reactor cattle (10). Only the results of the pool/antigen stimulating the greatest IFN-γ response are shown. Green squares represent *M. bovis* infected cattle, red triangles represent PP-A reactors and blue circles represent BCG vaccinated cattle. Dashed horizontal line indicate the positive cut0off ($OD_{450}$ values with antigens minus $OD_{450}$ without antigens $\geq$0.1).

Figure 3 shows IFN-γ secretion induced by individual peptides from pool 3 (A) and pool 26 (B) in whole blood cultures from two representative animals. Whole blood was collected from *M. bovis* experimentally infected cattle and incubated for 48hrs with peptides (25ug/ml each). Results are expressed as delta mean optical density $OD_{450}$ values with antigens minus $OD_{450}$ without antigens) of duplicate determinations, with a positive cut-off of 0.1.

Figure 4 shows the antigens selected for evaluation.

Figure 5 shows the most frequently recognised antigen.

Figure 6 shows the sequence homology between peptide 3.2 from Rv3873 (shown as SEQ ID NO. 7) with other mycobacterial proteins.

Figure 7 shows the amino acid sequences which refer to the open reading frames Rv1979c, Rv1769c, Rv1986, Rv3872, Rv3878, Rv1983, Rv3873 and Rv3879c which are shown as SEQ ID Nos 1 to 6.

Figure 8 shows the nucleotide sequences of the Rv1979c, Rv1769c, Rv1986, Rv3872, Rv3878, Rv1983, Rv3873 and Rv3879c antigens which are shown as SEQ ID Nos 8 to 13.

The following results demonstrate that six antigens showed promise as diagnostic antigens with regard to their specificity, and that two more could be considered as potential vaccine candidates because they were highly immunogenic in all groups assayed.

## MATERIALS AND METHODS

**Cattle.** Ca. 6 months old calves (Friesian or Friesian crosses) were obtained from herds free of bovine tuberculosis.

The following groups of cattle were used in this study:

***M. bovis infection***. Nine calves were infected with a GB *M. bovis* field strain from (AF 2122/97) by intratracheal instillation of $2 \times 10^4$ CFU as described [Buddle, 1995 #16; Rhodes SG, 2000 #138]. Twelve calves were infected with an *M. bovis* field strain, isolated from a New Zealand infected cow using also intratracheal instillation ($5 \times 10^3$ CFU). Bovine tuberculosis was confirmed in these animals by the presence of visible lessions in lymph nodes and lungs found at post-mortem examinations, by the histo-pathological examination of lesioned tissues and the culture of M. bovis from tissue samples collected from lymph nodes and lungs. Heparinised blood samples were obtained between 14-20 weeks after infection when strong and sustained *in vitro* tuberculin responses were observed. Data from a total of 21 experimentally infected cattle are presented in this study. One naturally infected animal was also used included in this group.

***BCG vaccination.*** Calves were vaccinated with BCG Pasteur by subcutaneous injection of $10^6$ CFU into the side of the neck followed 8 weeks later by a booster injection using the same route and dose [Buddle, 1995 #16; Vordermeier, 1999 #125]. Heparinised blood samples were taken between 4-6 weeks after the booster vaccination. Data from 6 calves will be presented in this study.

***Uninfected controls.*** Heparinised blood from tuberculin skin test-negative calves from herds free of BTB (10 animals) was also obtained. These animals produced IFN-γ *in vitro* after stimulation with tuberculin from *M. avium* indicating that they have been exposed to environmental mycobacteria.

***Antigens and peptides***

Antigens: Bovine (PPD-B) and avian (PPD-A) tuberculins were obtained from the Tuberculin Production Units at the Veterinary Laboratories Agency-Weybridge and used in culture at 10 μg/ml.

Peptides: A set of five hundred and fifty two synthetic peptides spanning 13 open reading frames (20 residues long with a 12 residue overlap) was prepared by Multi-rod peptide synthesis. These were used in mapping experiments in pools of 10 peptides at 5μg each peptide/ml and 25μg/ml when used individually. The peptides were purchased from Chiron Mimotopes (Clayton, Australia). ESAT-6 and CFP-10 derived peptides were synthesised by solid phase peptide synthesis and formulated into a peptide cocktail as described earlier [Vordermeier, 2001 #169]. They were also used at 5μg each peptide/ml. Peptide purity and sequence fidelity of ESAT-6 and CFP-10 derived peptides was confirmed by analytical reverse-phase HPLC and by electron-spray mass spectrometry, respectively.

***Interferon-gammaELISA.*** Whole blood cultures were performed in 96-well plates in 0.2ml/well aliquots by mixing 0.1 ml of heparinised blood with an equal volume of antigen containing solution [Vordermeier, 1999 #125]. Supernatants were harvested after 48 h of culture at 37°C/5% $CO_2$ in a humidified incubator. Interferon-gamma (IFN-γ) concentration was determined using BOVIGAM™ ELISA kit (Biocore AH, Omaha, NE). Results were deemed positive when the $OD_{450}$ [PPD-B] minus $OD_{450}$

values with antigens minus $OD_{450}$ value without antigens were $\geq$ 0.1. For comparative analysis of PPD-B vs. PPD-A responses, a positive result was defined by an $OD_{450}$ [PPD-A $\geq$0.1, and $OD_{450}$ [PPD-B] minus $OD_{450}$ [unstimulated] $\geq$0.1.

***Bioinformatics***

The DNA sequence of *M. tuberculosis* H37Rv was visualised using either the ARTEMIS display tool [Rutherford, 2000 #183] or the TubercuList database (http://genolist.pasteur.fr/TubercuList/) BLAST searches were performed from within TubercuList, or using the NCBI BLAST server (http://www.ncbi/nlm.nih.gov/BLAST)

## RESULTS

### EXAMPLE 1

**Selection of candidate antigens from the RD1, RD2, and RD14 regions of *M. bovis*.**

Thirteen ORFs from the RD1, RD2 and RD14 regions of *M. bovis* were selected for screening. These regions are deleted in BCG Pasteur and proteins encoded within these regions hold promise as candidate antigens for the differential diagnosis of *M. bovis* infected animals from BCG vaccinated cattle and as potential vaccine candidates. Selection criteria were that the ORF should encode a protein that either (i) showed no, or minimal, sequence similarity to other proteins in *M. tuberculosis* or other organisms, (ii) belonged to the PE or PPE protein family, (iii) had the potential of being induced or upregulated *in vivo* (e.g. amino acid transporters), or (iv) had the potential to be secreted. The designations of the antigens encoded by the selected ORF (Rv number), their sizes, and putative functions are listed in Figure 4.

**Immunogenicity of selected antigens in *M. bovis* infected, BCG vaccinated and environmentally sensitised cattle.**

Five hundred and thirty six overlapping peptides derived from the sequences of these antigens were synthesised. Peptides were then formulated into pools of approximately 10 neighbouring overlapping peptides, which resulted in 52 peptide pools. Figure 4 indicates the pool in relation to the antigens they represent as well as the total number of peptides/antigen required to ensure complete sequence coverage. Blood samples were obtained from 22 *M. bovis* infected animals, 6 *m. bovis* BCG

Pasteur vaccinated animals and 10 un-vaccinated/un-infected controls. Whole blood cultures in the presence of PPD-B, PPD-A, peptide pools and a cocktail of 10 synthetic peptides derived from ESAT-6 and CFP-10, were established and the amount of IFN-γ determined after 48 h of culture.

As expected, all *M. bovis* infected and BCG vaccinated animals responded more strongly to bovine tuberculin PPD-B than to avian tuberculin PPD-A (median responses and range: *M. bovis* infected: PPD-B=1.593(0.274-3.500), PPD-A=1.313(0.066-3.455)); BCG vaccinated: PPD-B=0.886(0.181-2.244), PPD-A=0.5115(0.274-2.234)); Uninfected, non-vaccinated control against animals responded strongly to avian PPD (PPD-A) indicating that they were sensitised by environmental mycobacteria (Median responses and ranges: PPDB=0.230(0.090-0.684), PPD-A=0.686(0.162-1.822)); they will be described hereinafter as *PPD-A reactors*. Next the immunogenicity of the peptide pools described in Figure 4 was assessed. Figure 1 depicts the results obtained with blood from two representative animals, one infected with *M. bovis*, the other a PPD-A reactor. The *M. bovis* infected animal recognised at least one peptide pool from each antigen (Figure 1A, C, E), indicating that cellular responses were induced after *M. bovis* infection against all 13 antigens selected. In contrast, none of the peptide pools induced IFN-γ secretion in whole blood from the environmentally sensitised PPD-A reactor (Figure 1B, D F).

The peptide-induced IFN-γ responses of all 38 M. bovis infected, BCG vaccinated and PPD-A-reactors (uninfected controls) to the 13 antigens are summarised in Figure 2. When antigens were covered by more than one peptide pool, the result of the pool stimulating the most IFN-γ secretion is shown. Interestingly, all 13 antigens were recognised by *M. bovis* infected cattle all be it with the percentage of responding cattle (responder frequencies) varying between 21 and 86 %. The most frequently recognised antigens were Rv3873, Rv3879c and Rv1769, with responder frequencies of 82, 77 and 86% respectively, whereas Rv1984c and Rv1772 were recognised only by 21 and 36% of infected calves. Interestingly, several of the most prominently recognised antigens were members of the PE/PPE protein family (e.g., Rv3873, with a responder frequency of 82%). Surprisingly, considering the absence of the genes encoding these antigens in BCG Pasteur, 9 of the 13 antigens tested, stimulated a positive response in BCG vaccinated animals (Rv3873, Rv3879c, Rv1979c, Rv1983,

Rv1987, Rv1989c, Rv1768, Rv1769 and RV 1772, with a range in responder frequencies of 17 – 100%). The remaining four antigens were recognised by *M. bovis* infected cattle only (Rv3872, Rv3878, Rv1984c, and Rv1986, with a range of responder frequencies of 21-59%). The responder frequencies of the 8 most immunogenic antigens are summarised in Figure 5. In addition, 21/22 *M. bovis* infected animals responded to a previously characterised peptide cocktail derived from CFP-10 and ESAT-6 (reference) that had been included for comparison (median responses and range: 1.281 (0.011-2.825)).

**EXAMPLE 2**

**The combination of antigens offers improved sensitivity**

It is unlikely that a single diagnostic antigen, however specific, could impart enough sensitivity to provide population coverage; therefore combinations of specific antigens will be needed. It was therefore determined whether such antigen combinations could improve test sensitivity. Two scenarios were considered: firstly, antigens suitable for differential diagnosis, i.e., not recognised by BCG vaccinated animals or PPD-A reactors. The three antigens most frequently recognised by *M. bovis* infected animals fulfilling this criteria are Rv1986, Rv3872, and Rv3878 (Figure 5). Combining their results indicated that 82% of the infected animals would have been correctly identified by their responses to either of these three antigens (Figure 5).

Secondly, we considered the three most immunodominant antigens (Rv1983, Rv3873, Rv3879c) that were not recognised by PPD-A reactors, but were recognised by BCG vaccinated calves, i.e., antigens capable of distinguishing between *M. bovis* infection and animals sensitised by environmental mycobacteria for example, *M. avium* (specific diagnosis). Taken together, these antigens would have identified 20/22 (91%) of the *M. bovis* infected animals (Figure 5). Interestingly, if Rv3878 from the first category was considered together with Rv3873 and Rv3879c from this category. 21/22 (95%) of the *M. bovis* infected animals would have been detected (Figure 5).

**EXAMPLE 3**

**Responses of peptide pools can be the result of a single peptide**

The peptide pools formulated contain between 8-11 peptides (see Figure 4 for details of peptide pools). To determine whether IFN-γ responses of pools were due to single or multiple peptide constituents, the individual peptides of pool 3 (representing residues 89-188 from Rv3873) and pool 26 (representing residues 161-252 from Rv1983) were tested using blood from *5 M. bovis* infected animals. All three animals tested that recognised pool 3 responded exclusively to peptide 3.2 (residues 97-116), whereas both animals tested that responded to pool 26 only recognised peptide 26.2 (residues 169-188). The results shown in Figure 3 give results from one representative animal responding to pools 3 (Fig. 3A) or 26 (Fig.3B), respectively. These data suggest that the individual peptides imparting antigenicity can be identified from immunodominant pools and that pool immunogenicity can be attributed to single peptides.

The effective use of comparative genomics in combination with synthetic peptides to identify and screen thirteen potential antigens encoded by ORFs located in the RD1, RD2, and RD14 regions of the *M. tuberculosis* has been demonstrated. These results indicated that six antigens in particular showed promise as diagnostic antigens because they were either (i) recognised by *M. bovis* infected animals alone, but not by BCG vaccinated or controls (differential diagnosis, Figure 5) or (ii) by infected animals and vaccinated animals but not by environmental mycobacteria exposed controls (specific diagnosis, Figure 5).

In general, all 13 antigens tested were recognised with responder frequencies varying between 21 and 86%. It is likely that a combination of several factors determines whether and to what degree mycobacterial proteins are immunogenic after infection. These factors could include (a) parameters intrinsic to the bacterium, such as the abundance of the protein, its sub-cellular location, post-translational modification, participation in macromolecular complexes, and *in vivo* regulation; and (b) factors relating to the immune system, including location of the antigen with respect to the phagosome, proteolytic sensitivity, and the presence of motifs suitable for interaction with TAP transporters and different MHC alleles within the antigen.

The present invention exploits the use of pools of overlapping synthetic peptides derived from the sequences of these proteins. In a pilot experiment where the

peripheral monocyte blood cell (PBMC) was isolated from 8 cattle experimentally infected with *M. bovis* and stimulated them with either recombinant ESAT-6 or a cocktail of 11 synthetic peptides spanning the whole sequence of ESAT-6, it was concluded that the numbers of IFN-γ producing cells, determined in this case by ELISPOT, demonstrated equivalent responses to recombinant protein and synthetic peptides ($r=0.92$, $p<0.0001$.) The number of peptide pools that represent the sequences of each ORF varies depending on the size of the antigen, as illustrated in Figure 4. It was demonstrated that the combined results from Rv3873, Rv3878 and Rv3879c resulted in an overall responder frequency of 95%. These 3 antigens are represented by a total of 16 different peptide pools, containing 169 individual peptides. However, the same frequency of recognition can be obtained using just 3 pools out of the 16 pools assayed (pools 3, 8 and 9), i.e., 30 peptides, suggesting the presence of the immunodominant epitomes within these three pools. Indeed, the number of peptides needed to achieve responder frequencies similar to that with the complete set of overlapping peptides could even be significantly lower since the data described in Figure 3 demonstrates that only one or two immunodominant peptides can be responsible for the immunogenicity of the whole pool. If these peptides were to be recognised promiscuously in the context of multiple MHC molecules, as has been described in the recognition of other mycobacterial antigens by human, murine and bovine CD4+ T cells [Vordermeier, 2000 #126; Vordermeier, 2001 #169; Vordermeier, 1995 #124; Lightbody, 1998 #75; Lightbody, 1998 #76; Pollock, 1994 #99; Pollock, 1995 #100], the number of peptides required to achieve wide population coverage could be relatively low as has been demonstrated before for ESAT-6 and CFP-10 derived peptides [Vordermeier, 2001 #169, Lalvani, 2001 #185]. Peptides as immuno-diagnostic reagents can therefore constitute a practical alternative to recombinant proteins, in addition to substituting them as reagents to assess immunogenicity. The fact that all three animals tested, two from the UK and one from New Zealand, recognised the same peptide within pool 3 (peptide 3.2) is encouraging in this context.

Interestingly, the previously described peptide cocktail containing peptides derived from ESAT-6 and CFP-10 was also recognised by 95% of the *M. bovis* infected animals tested, in fact the same animals that responded to the combination of Rv3873, Rv3878 and Rv3879c.

As described in Figure 6, 4 PPD/ PE genes were selected for testing (Rv3872, Rv3873, Rv1983 and Rv1768) and gave responder frequencies of between 45-82% when assayed in the *M. bovis* infected cattle. Little is known about the function or immunogenicity of these proteins, which account for approximately 10% of the total coding capacity of the *M. tuberculosis* genome [Cole, 1998 #21].

As described in Figure 3, peptide 3.2 is a highly immunogenic component of pool 3 derived from the sequence of Rv3873, a member of the PPE family of proteins. The pool consistently produces positive responses when assayed in *M. bovis* infected cattle with a responder frequency of 82% but was also recognised in BCG vaccinated animals. This is a surprising outcome given that its gene is deleted in BCG and that no homologous proteins were found elsewhere in the BCG genome. However, the unit of cross-reactivity is the epitope, less than 20 amino acids long, that is recognised by T cells in the context of MHC molecules [Rudensky, 1991 #175, Rammensee, 1995 #176]. Consequently, the molecular nature of cross-reactivity can only be addressed once these epitopes have been identified [Harris, 1995 #177]. Therefore we used the sequence of peptide 3.2 (shown as SEQ ID NO.9) to search for similar regions with other genes found within the *M. tuberculosis* genome.

Figure 6 shows the results using the Basic Local Alignment Search Tool (BLAST) program [NCBI, #178] to identify similarity between mycobacterial proteins. The table shown in Figure 6 highlights several sequences that contain amino acid identities of greater than 50%. These include five proteins from the *M. tuberculosis* genome, all of which are also members of the PPE family and several others identified in proteins of various mycobacterial species. The peptide covers an area of the gene that encodes two motifs identified in a number of PPE family members during their annotation [Tekaia, 1999 #116, TubercuList, #180]. Therefore, it is reasonable to hypothesise that the cross reactive nature of the peptide is a result of similarity with other PPE family members located elsewhere in the genome of *M. tuberculosis* and therefore the genome of *M. bovis* BCG Pasteur. We conducted BLAST searches for the other identified cross-reactive antigens (e.g., Rv1979c) by comparing the whole genes in steps of 20 amino acids, representing the corresponding peptides, and were able to

find numerous similar amino acid sequences in other mycobacterial proteins outside the deleted regions.

The use of peptides instead of recombinant proteins, has several advantages already discussed. However, with regard to the observed cross reactivity of antigens between BCG vaccinated and *M. bovis* infected animals, this peptide-based approach has other distinct advantages. If ORF Rv1987 is taken as an example, it appears unsuitable as a differential diagnostic reagent due to the high cross reactivity in the BCG vaccinated cattle. However, the responder frequency of 57% in *M. bovis* infected cattle is due to the recognition of two pools with responder frequencies of 47% and 53% respectively. Whilst one pool is recognised by 50% of the BCG vaccinated animals, the other is not recognised. Therefore, the diagnostic potentials of this antigen can still be realised by using only peptides derived from the second peptide pool.

In summary, therefore, the analysis of peptides, derived from genes deleted in BCG Pasteur, can lead to the identification of antigens for diagnosis and even vaccination. In particular, antigens that can form the basis of diagnostic reagents to either differentiate between infected and BCG vaccinated animals or to improve the specificity of PPD *per se* are described. In addition, it has also been demonstrated for the first time that members of the both the PE and PPE families of proteins induced cellular immune response after mycobacterial infection of a target species.

All references mentioned in the above specification are herein incorporated by reference. Other modifications of the present invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with the specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention, which are obvious to those skilled in the art, are intended to be within the scope of the following claims.

*REFERENCES*

1. **Anderson, P.** 2001, TB Vaccines: progress and problems. Trends Immunol. **22**:160-8.2.

2. **Arend, S M., A. Geluk, K. E. Van Meijgaarden, J. T. Van Dissel, M. Theisen, P. Anderson, and T. H. M. Ottenhoff.** 2000. Antigenic equivalence of human T cell responses to Mycobacterium tuberculosis specific RD1- encoded protein antigens ESAT-6 and culture Filtrate Protein 10 and to mixtures of synthetic peptides. Infect. Immun. **68**:3314-3321.

3. **Behr, M. A., M. A. Wilson, W. P. Gill, H. Salmon, G. K. Schoolnik, S. Rane, and P. M. Small.** 1999. Comparative genomics of BCG vaccines by whole genome DNA microarray. Science **284**:1520 - 1523.

4. **Berggren, S., A.** 1981. Field experiment with BCG vaccine in Malawi. Br. Vet. J. **137**:88 - 94.

5. **Buddle, B., N. A. Parlane, D. L. Keen, F. E. Aldwell, J. M. Pollock, K Lightbody, and P. Andersen.** 1999. Differentiation between *Mycobacterium bovis* BCG-vaccinated and *M. bovis* infected cattle using recombinant mycobacterial antigens. Clin. And Diagn. Lab. Immunol. **6**:1-5.

6. **Buddle, B., B Ward, F. Aldwell, D. Collins, and G. de Lisle.** 2002. Influence of sensitisation to environmental mycobacteria on subsequent vaccination against bovine tuberculosis. Vaccine **20**: 1126-33.

7. **Buddle, B., G. W. de Lisle, A. Pfeffer, and F. E. Aldwell.** 1995. Immunological responses and protection against Mycobacterium bovis in calves vaccinated with a low dose of BCG. Vaccine **13**: 1123-30.

8. **Buddle, B., D. Keen, A. Thomson, G. Jowett, A. R. McCarthy, J. Heslop, G. W. de Lisle, J. L. Stanford, and F. E. Aldwell.** 1995. Protection of cattle

from bovine tuberculosis by vaccination with BCG by the respiratory or subcutaneous route, but not by vaccination with killed *Mycobacterium vaccae*. Res. Vet. Sci. **59**: 10-6.

9. **Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry, 3rd, Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Delvin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, B. G. Barrell, and et al.** 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature **393**: 537-44.

10. **Cooper, A., D. Dalton, T. Stewart, J. Griffin, D. Russell, and I. Orme.** 1993. Disseminated tuberculosis in interferon gamma gene-disrupted mice. J. Exp. Med. **178**: 2243-47.

11. **Cosivi, O., F. X. Meslin, C. J. Daborn, and J. M. Grange.** 1995. Epidemiology of *Mycobacterium bovis* infections in animals and humans, with particular reference to Africa. Rev. Sci, Tech. **14**:733-46.

12. **Delogu, G., and M. J. Brennan.** 2001. Comparative immune response to PE and PE_PGRS amtigens of *Mycobacterium tuberculosis*. Infect. Immun. **69**:5606-11.

13. **Fine, P. E.** 1989. The BCG story: lessons from the past and implications for the future. Rev. Infect. Dis. **11** Suppl 2: S353-9.

14. **Flynn, J., J. Chan, K. Triebold, D. Dalton, T. Stewart, and B. Bloom.** 1993. An essential role for interferon gamma in resistance to *Mycobacterium tuberculosis* infection. J. Exp. Med. **178**:2249-54.

15. **Gordon, S. V., R. Brosch, A. Billault, T. Garnier, K. Eiglmeier, and S. T. Cole.** 1999. Identification of variable regions in the genomes of tubercle

bacilli using bacterial artificial chromosome arrays. Mol. Microbiol. **32**: 643-655.

*16.* **Hardie, R. M., and J. M. Watson.** 1992. *Mycobacterium bovis* in England and Wales: past, present and future. Epidemiol. Infect. **109**: 23-33.

*17.* **Harris, D., H. Vordermeier, M. Singh, C. Moreno, S. Jurcevic, and J. Ivanyi.** 1995. Cross-recognition by T cells of an epitope shared by two unrelated mycobacterial antigens. Eur. J. Immunol. **25**:3173-9.

*18.* **Hubrig, T., and W. Kruger.** 1958. Untersuchungen zur Tuberkuloseschutzimpfung bei Rindern. Monatshefte fuer Veterinaermedizin **13**:513 – 519.

*19.* **Jouanguy, E., F. Altare, S. Lamhamedi, P. Revy, J. F. Emile, M. Newport, M. Levin, S. Blanche, E. Seboun, A. Fischer, and J. L. Casanova.** 1996. Interferon-gamma-receptor deficiency in an infant with fatal bacille Calmette-Guerin infection. N. Engl. J. Med. **335**:1956-61.

*20.* **Krebs, J. R.** 1997. Bovine tuberculosis in cattle and badgers. Ministry of Agriculture, Fisheries and Food Publications, London UK.

*21.* **Lalvani, A., P. Nagvenkar, Z. Udwadia, A. Pathan, K. Wilkinson, J. Shastri, K. Ewer, A. Hill, A. Mehta, and C. Rodrigues.** 2001. Enumeration of T cells specific for RD1-encoded antigens suggests a high prevalence of latent *Mycobacterium tuberculosis* infection in healthy urban Indians. J. Infet. Dis. **183**:469-477.

*22.* **Lalvani, A., A. Pathan, H. McShane, R. Wilkinson, M. Latif, C. Conlon, G. Pasvol, and A. Hill.** 2001. Rapid detection of *Mycobacterium tuberculosis* infection by enumeration of antigen-specific T cells. Am. J. Respir. Crit. Care. Med. **163**: 824-8.

23. **Lightbody, K. A., R. M. Girvin, D. P. Mackie, S. D. Neill, and J. M. Pollock.** 1998. T-cell recognition of mycobacterial proteins MPB70 and MPB64 in cattle immunized with antigen and infected with *Mycobacterium bovis*. Scand. J. Immunol. **48**: 44-51.

24. **Lightbody, K. A., R. M. Girvin, D. A. Pollock, D. P. Mackie, S. D. Neill, amd J. M. Pollock.** 1998. Recognition of a common mycobacterial T-cell epitope in MPB59 of *Mycobacterium bovis*. Immunology **93**:314-22.

25. **Mahairas, G. G., P. J. Sabo, M. J. Hickey, D. C. Singh, and C. K. Stover.** 1996. Molecular analysis of genetic differences between *Mycobacterium bovis* BCG and virulent *M. Bovis* J. Bacteriol. **178**:1274-82.

26. **NCBI.** Basic Local Alignment Search Tool (BLAST). http://www/ncbi.nlm.nih.gov/BLAST/.

27. **Pollock J. M., and P. Andersen.** 1997. The potential of the ESAT-6 antigen secreted by virulent mycobacteria for specific diagnosis of tuberculosis. J. Infect. Dis. **175**:1251-4.

28. **Pollock, J. M., A. J. Douglas, D. P. Mackie, and S. D. Neill.** 1994. Identification of bovine T-cell epitopes for three *Mycobacterium bovis* antigens: MPB70, 19,000 MW and MPB57. Immunology. **82**:9-15.

29. **Pollock, J. M., A. J. Douglas, D. P. Mackie, and S. D. Neill.** 1995. Peptide mapping of bovine T-cell epitopes for the 38 kDa tuberculosis antigen. Scand. J. Immunol. **41**:85-93.

30. **Rhodes SG, Gavier-Widen D, Buddle, BM., Whelan AO, Singh M, Hewinson RG, and V. HM.** 2000. Antigen specificity in experimental bovine tuberculosis. Infect. Immun **68**:2573-2578.

31. **Rutherford, K., J. Parkhill, J. Crook, T. Hornsell, P. Rice, M. Rajandream, and B. Barrell.** 2000. Artemis: Sequence visualisation and annotation. Bioinformatics B10: 944 – 5.

32. **Tekaia, F., S. V. Gordon, T. Garnier, R. Brosch, B. G. Barrell, and S. T. Cole.** 1999. Analysis of the proteome of *Mycobacterium tuberculosis* in silico. Tuber. Lung Dis. **79**:329-42.

33. TubercuList. MAST – Motif Alignment and Search Tool http://genolist.pastuer.fr/TubercuList/.

34. **Van pinxteren, L. A., P. Ravn, E. M. Agger, J. Pollock, and P. Andersen.** 2000. Diagnosis of Tuberculosis Based on the two specific Antigens ESAT-6 and CFP10. Clin. Diagn. Lab. Immunol. **7**:155-160.

35. **Vordermeier, H. M.** 1995. T-cell recognition of mycobacterial antigens. Eur. Respir. J. Suppl. **20**:657s-667s.

36. **Vordermeier, H. M., P. C. Cockle, A. Whelan, S. Rhodes, and G. Hewinson.** 2000. Towards development of diagnostic assays to discriminate between *Mycobacterium bovis* infection and BCG vaccination in cattle. Clin. Infect. Dis. **30**:S291-S298.

37. **Vordermeier, H. M., P. C. Cockle, A. Whelan, S. Rhodes, N. Palmer, D. Bakker, and R. G. Hewinson.** 1999. Development of diagnostic reagents to differentiate between *Mycobacterium bovis* BCG vaccination and *M. bovis* infection in cattle. Clin. Diagn. Lab. Immunol. **6**:675-682.

38. **Vordermeier, H. M., A. Whelan, P. C. Cockle, L. Farrant, N. Palmer, and R. G. Hewinson.** 2001. Use of synthetic peptides derived from the antigens ESAT-6 and CFP-10 for differential diagnosis of bovine tuberculosis in cattle. Clin. Diagn. Lab. Immunol. **8**:571-8.

39. **Wood, P. R., and J. S. Rothel.** 1994. In vitro immunodiagnostic assays for bovine tuberculosis. Vet. Microbiol. **40**:125-135.

## CLAIMS

1. A diagnostic reagent comprising a polypeptide derived from an RD1 or RD2 regions of the *Mycobacterium tuberculosis, Mycobacterium bovis* or *Mycobacterium africanum* genomes, or a variant, homologue or fragment of these, with the proviso that the polypeptide is not a ESAT-6, CFP-10, MPT-64 or a polypeptide encoded by the Rv1984c, Rv3871, Rv3872 or Rv3873 regions of the *Mycobacterium tuberculosis, Mycobacterium bovis* or *Mycobacterium africanum* genomes.

2. A diagnostic reagent according to claim 1 wherein the polypeptide is derived from the *Mycobacterium tuberculosis* genome.

3. A diagnostic reagent according to claim 1 which polypeptide is a member of the PE/PPE protein family.

4. A diagnostic reagent according to any one of claims 1 to 3 wherein the polypeptide comprises the sequence shown in any one of SEQ ID Nos 1, 3, 4 or 6, or a fragment, homologue or variant thereof.

5. A diagnostic reagent according to claim 4 wherein the diagnostic reagent comprises the sequence shown in SEQ ID Nos 1 or 3 or a fragment, homologue or variant thereof.

6. A diagnostic reagent according to any one of the preceding claims wherein the diagnostic reagent is used to differentiate between tuberculosis-infected and tuberculosis-vaccinated mammals.

7. A diagnostic reagent according to claim 6 wherein the diagnostic reagent is used to differentiate between *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobacterium africanum*-infected and mammals vaccinated against *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobacterium africanum*..

8. A diagnostic reagent according to any one of claims 1 to 4 wherein the diagnostic reagent comprises the sequence shown in SEQ ID Nos 3 or 6 or a fragment, homologue or variant thereof.

9. A diagnostic reagent according to claim 8 wherein the diagnostic reagent is used to differentiate between those mammals which are either vaccinated against or infected by tuberculosis and those mammals sensitised by environmental mycobacteria.

10. A diagnostic reagent according to claim 9 wherein the diagnostic reagent is used to differentiate between those mammals which are either vaccinated against or infected by *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobacterium africanum* and those mammals sensitised by environmental mycobacteria.

11. A peptide derived from an RD1 or RD2 region of the *Mycobacterium tuberculosis, Mycobacterium bovis* or *Mycobacterium africanum* genomes, or a variant, homologue or fragment of these, with the proviso that the peptide is not derived from a ESAT-6 or CFP-10 polypeptide.

12. A peptide according to claim 11 wherein the peptide has the sequence shown in SEQ ID NO. 7, or a fragment, homologue or variant thereof.

13. A peptide according to claim 11 or claim 12 wherein the peptide is a synthetic peptide.

14. A peptide according to any one of claims 11 to 13 which is a diagnostic reagent.

15. A peptide according to claim 14 wherein said peptide is used to differentiate between mammals which are either vaccinated against or infected by *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobacterium africanum* and mammals sensitised by environmental mycobacteria.

16. A peptide according to claim 15 wherein said peptide is used to differentiate between those mammals infected by *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobacterium africanum* and those mammals sensitised by environmental mycobacteria.

17. A nucleic acid encoding the polypeptide or peptide sequence of any one of the preceding claims or a fragment, homologue or variant thereof.

18. A nucleic acid according to claim 17 wherein the nucleic acid comprises the sequence shown in SEQ ID Nos 8, 10, 11 or 13, or a variant or fragment thereof.

19. A diagnostic kit comprising at least one of the polypeptides or peptides encoded by the sequences shown as SEQ ID Nos 1, 3, 4, 6 and 7, and optionally at least one polypeptide encoded by the sequences shown as SEQ ID Nos 2 and 5, and optionally one or more reagents, wherein the polypeptides or peptides are capable of differentiating between tuberculosis-infected, tuberculosis-vaccinated mammals and mammals which have been sensitised by environmental bacteria.

20. A diagnostic kit according to claim 19 wherein the kit comprises the polypeptides encoded by the sequences shown as SEQ ID Nos 1, 2 and 3.

21. A diagnostic kit according to claim 19 or claim 20 wherein the diagnostic kit differentiates between those mammals infected by *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobacterium africanum* and those mammals vaccinated against *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobacterium africanum* .

22. A diagnostic kit according to claim 19 wherein the kit comprises the polypeptides or peptides encoded by the sequences shown in SEQ ID Nos 4, 5 and 6 and optionally 7.

23. A diagnostic kit according to claim 22 wherein the diagnostic kit differentiates between those mammals which are either vaccinated against or infected by *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobacterium africanum* and those mammals which have been sensitised by environmental mycobacteria.

24. A diagnostic kit according to any one of claims 19 to 23 further comprising one or more polypeptides or peptides from the RD1 region of the *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobacterium africanum* genomes.

25. A diagnostic kit according to claim 24 wherein the kit comprises a peptide mixture composed of one or more polypeptides from the RD1 region of the *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobacterium africanum* genomes.

26. A diagnostic kit according to claim 24 or claim 25 wherein the one or more polypeptides include the ESAT-6 and CFP-10 polypeptides.

27. A method of diagnosing infection in a host, or exposure of a host, to a mycobacterium comprising

i) contacting a population of cells from the host with a polypeptide derived from an RD1 or RD2 region of the *Mycobacterium tuberculosis, Mycobacterium bovis* or *Mycobacterium africanum* genomes, or a variant, homologue or fragment of these, which polypeptide may be used as a diagnostic reagent, with the proviso that the polypeptide is not a ESAT-6, CFP-10, MPT-64, or a polypeptide encoded by the Rv1984c, Rv3871, Rv3872 or Rv3873 regions of the *Mycobacterium tuberculosis, Mycobacterium bovis* or *Mycobacterium africanum* genomes; and

ii) determining whether the cells of said cell population recognise the polypeptide or fragment or variant thereof.

28. A method according to claim 27 wherein said cell population is a population of T-cells.

29. A method according to claim 27 or claim 28 wherein the infection is by *Mycobacterium tuberculosis, Mycobacterium bovis* or *Mycobacterium africanum.*

30. A pharmaceutical or veterinary composition comprising a polypeptide or peptide according to any one of claims 1 to 13, or a nucleic acid according to claim 17 or claim 18, in combination with a pharmaceutically or veterinarily acceptable carrier.

31. A polypeptide or peptide according to any one of claims 1 to 13, or a nucleic acid according to claim 17 or claim 18, for use as a medicament.

32. A polypeptide, peptide or nucleic acid according to claim 31 for use as a vaccine against tuberculosis caused by *Mycobacterium tuberculosis, Mycobacterium bovis* or *Mycobacterium africanum* in a mammal.

33. Use of a polypeptide or peptide according to any one of claims 1 to 13 in the preparation of a vaccine.

34. Use according to claim 33 wherein the vaccine is used to vaccinate against tuberculosis.

35. Use according to claim 33 or claim 34 wherein the vaccine is used to vaccinate against tuberculosis in cattle.

36. Use according to claim 33 or 34 wherein the vaccine is used to vaccinate against tuberculosis in humans.

37. Use according to any one of claims 33 to 36 wherein the vaccine comprises one or more subunits.

38. Use of a peptide according to any one of claims 11 to 13 to produce an antibody specific to the peptide.

39. A method of protecting a mammal against infection by *Mycobacterium bovis, Mycobacterium tuberculosis* or *Mycobacterium africanum* comprising administering to said mammal a polypeptide or peptide according to any one of claims 1 to 16, a nucleic acid according to claim 17, or a composition according to claim 28.

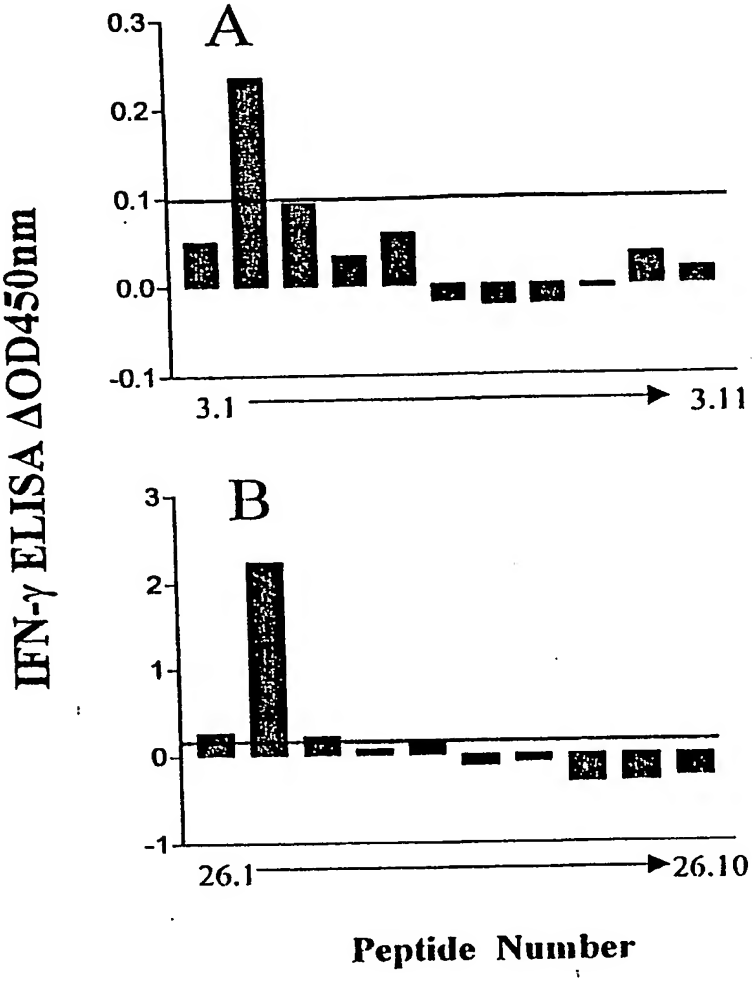
40. A polypeptide, a peptide, use or diagnostic kit as substantially as hereinbefore described with reference to the accompanying figures.

# ABSTRACT

## DIAGNOSTIC TEST

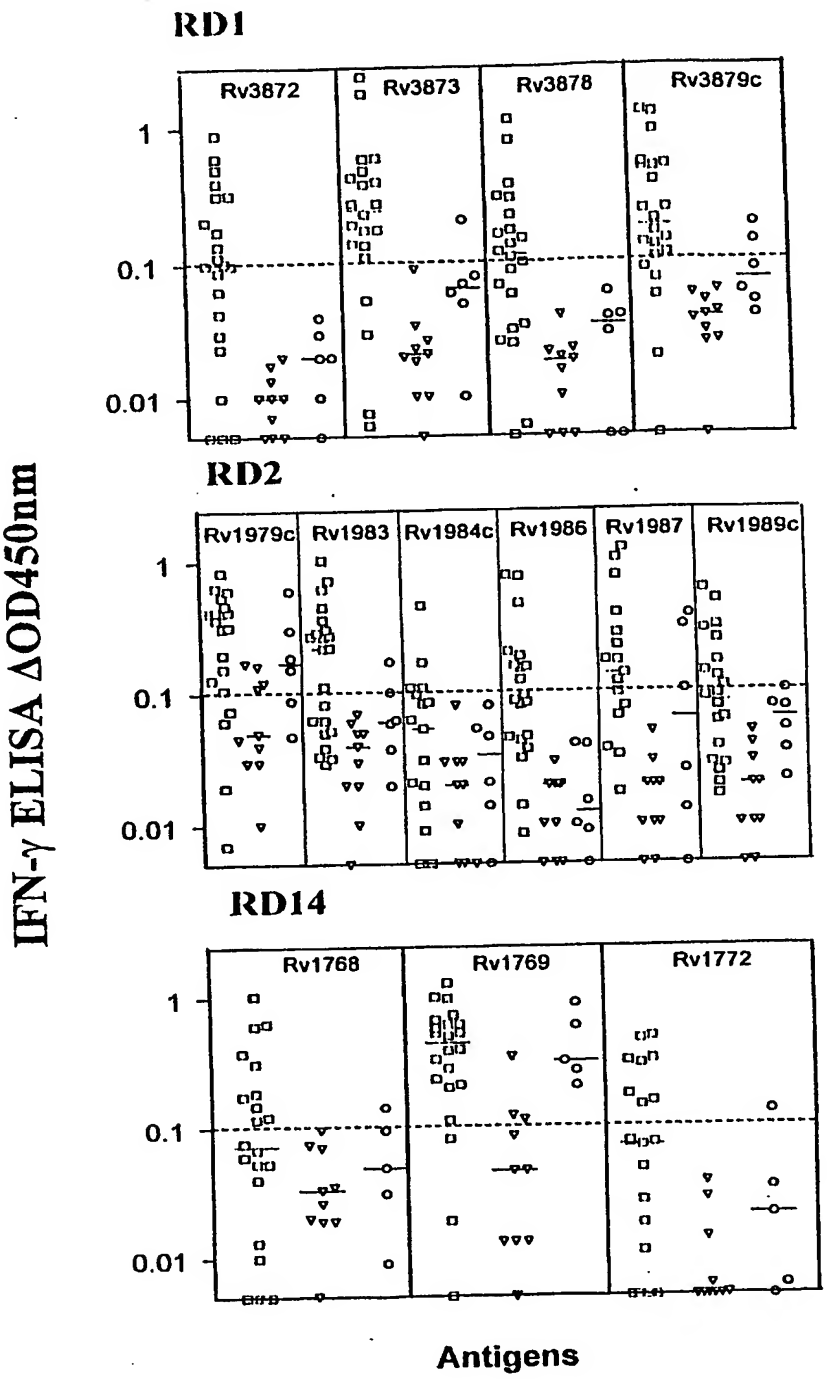
A diagnostic reagent comprising a polypeptide derived from the RD1 or RD2 regions of the *Mycobacterium tuberculosis,Mycobacterium bovis* or *Mycobacterium africanum* genomes, or a variant, homologue or fragment thereof, is described, with the proviso that the polypeptide is not the ESAT-6, CFP-10, MPT-64 or a polypeptide encoded by the Rv1984c, Rv3871, Rv3872 or Rv3873 regions of the *Mycobacterium tuberculosis, Mycobacterium bovis* or *Mycobacterium africanum* genomes. The diagnostic reagent may be used in a method for differentiating between mammals vaccinated against tuberculosis and tuberculosis-infected mammals and those mammals, which have been sensitised by environmental mycobacteria.

FIGURE 1



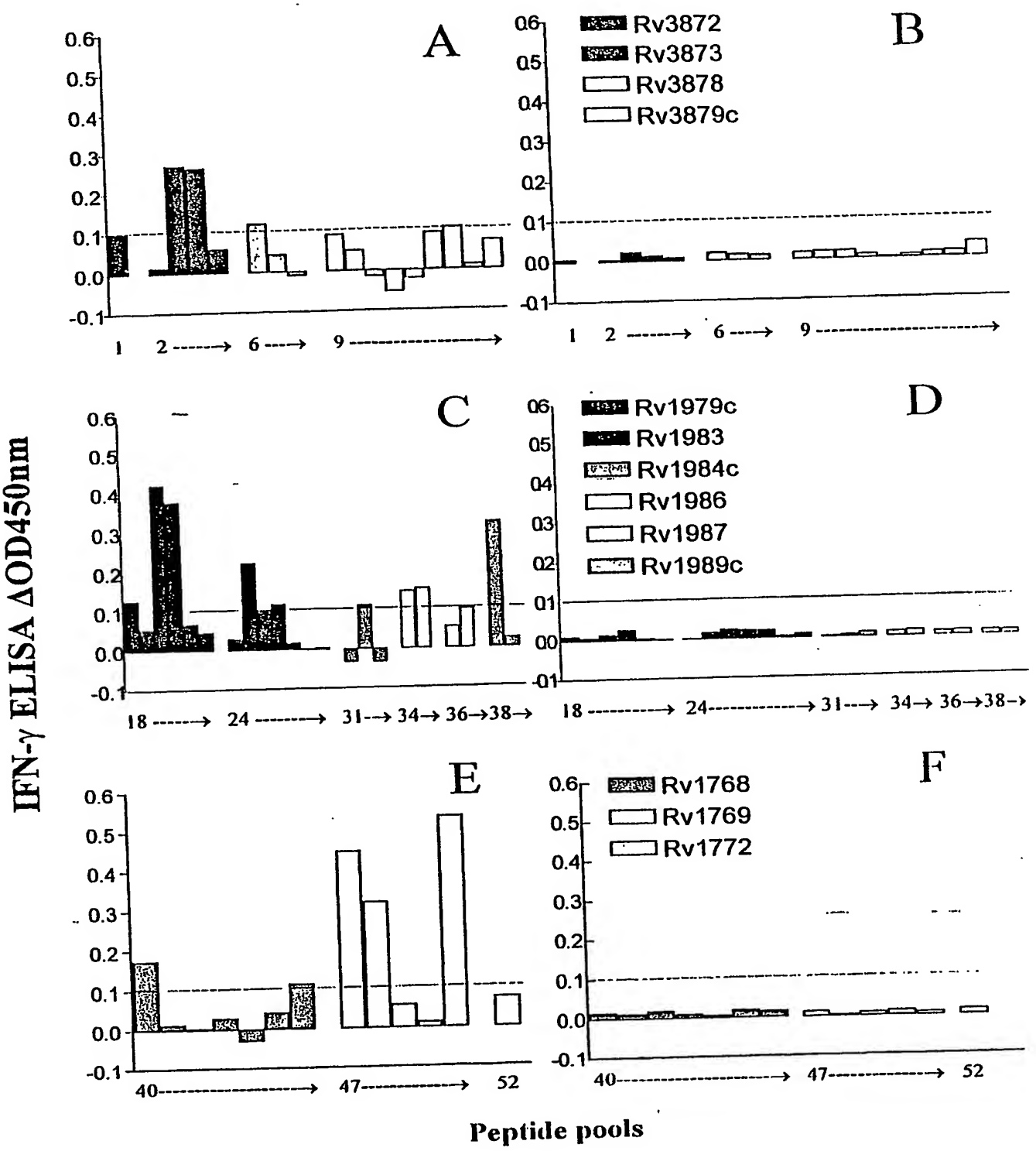

P.J. Cockle et al Fig. 3

FIGURE 2




P.J. Cockle et al Fig. 2

FIGURE 3




P.J. Cockle et al Fig. 1

Figure 4. RD antigens selected for evaluation in this study

| Deleted Region | Designation[a] | Size (Amino Acids) | Peptide Pools[b] | Putative Function[c] |
|---|---|---|---|---|
| RD1 | Rv3872 | 99 | 1 (10) | Member of PE-like protein family |
| | Rv3873 | 368 | 2-5 (40) | Member of *M. tuberculosis* PPE family |
| | Rv3878 | 280 | 6-8 (30) | Unknown, alanine-rich protein |
| | Rv3879c | 729 | 9-17 (90) | Unknown, alanine-proline-rich protein |
| RD2 | Rv1979c | 481 | 18-23 (60) | Possible amino acid permease |
| | Rv1983 | 558 | 24-30 (70) | Member of the PE-PGRS sub-family of glycine-rich proteins |
| | Rv1984c | 217 | 31-33 (30) | Probable secreted cutinase |
| | Rv1986 | 199 | 34-35 (20) | Possible lysine transporter |
| | Rv1987 | 142 | 36-37 (20) | Possible chitinase |
| | Rv1989c | 186 | 38-39 (20) | Unknown |
| RD14 | Rv1768 | 618 | 40-46 (70) | Member of the PE-PGRS sub-family of glycine-rich proteins |
| | Rv1769 | 414 | 47-51 (50) | Similar to *Streptomyces coelicolor* hypothetical protein |
| | Rv1772 | 103 | 51-52 (20) | Unknown |

[a]Rv designation of ORF as defined [Cole, 1998 #21]

FIGURE 4 (CONT)

[b]Number of peptide pools required to cover full sequence (total number of peptides required shown in brackets)

[c]Putative function as suggested [Cole, 1998 #21]

Figure 5. List of most frequently recognised antigens[a]

| Designation | Responder Frequency % | | | Potential Application |
|---|---|---|---|---|
| | [b]*M.bovis* Reactors | [c]BCG Vaccinated | [d]*M.avium* Reactors | |
| Rv1986 | 41 | 0 | 0 | Differential Diagnostics |
| Rv3872 | 50 | 0 | 0 | |
| Rv3878 | 59 | 0 | 0 | |
| Combined | 82 | 0 | 0 | |
| Rv1983 | 59 | 33 | 0 | Specific Diagnostics |
| Rv3873 | 82 | 17 | 0 | |
| Rv3879c | 77 | 33 | 0 | |
| Combined | 91 | 50 | 0 | |
| Rv1979c | 73 | 67 | 40 | Vaccines |
| Rv1769 | 86 | 100 | 30 | |

[a]Only antigens recognised by >40% of *M.bovis* infected animals are listed

[b]Results from 22 cattle experimentally infected with *M.bovis*

[c]Results from 5-6 BCG vaccinated cattle

[d]Results from 10 environmental mycobacteria sensitised cattle

Figure 6. Sequence homology between peptide 3.2 from Rv3873 with other mycobacterial proteins

| Designation[a] | Putative Function | Amino Acid Sequence[b] |
|---|---|---|
| **RV3873** | ***M.tuberculosis* PPE family** | **AMATTPSLPEIAANHIT** |
| Rv3021c | *M.tuberculosis* PPE family | ALAEMPTLPELAANHLT |
| Rv0286 | *M.tuberculosis* PPE family | ALAAMPTLAELAANHVI |
| Rv3018c | *M.tuberculosis* PPE family | ALAEMPTLPELAANHLT |
| Rv0280 | *M.tuberculosis* PPE family | AVAAMPTLVELAANHTL |

The homology search was performed using the BLAST program. [a]Designation of *M. tuberculosis* proteins as described [Cole, 1998 #21]. [b]The sequence in *M. tuberculosis* and *M. bovis* was found to be identical. Amino acid residues are shown in the one letter code. Non-identical residues are underlined.

FIGURE 7

>Rv1983: 558 aa - M. tuberculosis - SEQ.ID.No.4

```
  1 - VSFLVVVPEF LTSAAADVEN IGSTLRAANA AAAASTTALA AAGADEVSAA VAALFARFGQ
 61 - EYQAVSAQAS AFHQQFVQTL NSASGSYAAA EATIASQLQT AQHDLLGAVN APTETLLGRP
121 - LIGDGAPGTA TSPNGGAGGL LYGNGGNGYS ATASGVGGGA GGSAGLIGNG GAGGAGGPNA
181 - PGGAGGNGGW LLGNGGIGGP GGASSIPGMS GGAGGTGGAA GLLGWGANGG AGGLGDGVGV
241 - DRGTGGAGGR GGLLYGGYGV SGPGGDGRTV PLEIIHVTEP TVHANVNGGP TSTILVDTGS
301 - AGLVVSPEDV GGILGVLHMG LPTGLSISGY SGGLYYIFAT YTTTVDFGNG IVTAPTAVNV
361 - VLLSIPTSPF AISTYFSALL ADPTTTPFEA YFGAVGVDGV LGVGPNAVGP GPSIPTMALP
421 - GDLNQGVLID APAGELVFGP NPLPAPNVEV VGSPITTLYV KIDGGTPIPV PSIIDSGGVT
481 - GTIPSYVIGS GTLPANTNIE VYTSPGGDRL YAFNTNDYRP TVISSGLMNT GFLPFRFQPV
541 - YIDYSPSGIG TTVFDHPA
```

>Rv1986: 199 aa - M. tuberculosis - SEQ.ID.No.1

```
  1 - VNSPLVVGFL ACFTLIAAIG AQNAFVLRQG IQREHVLPVV ALCTVSDIVL IAAGIAGFGA
 61 - LIGAHPRALN VVKFGGAAFL IGYGLLAARR AWRPVALIPS GATPVRLAEV LVTCAAFTFL
121 - NPHVYLDTVV LLGALANEHS DQRWLFGLGA VTASAVWFAT LGFGAGRLRG LFTNPGSWRI
181 - LDGLIAVMMV ALGISLTVT
```

>Rv3872: 99 aa - M. tuberculosis - SEQ.ID.No.2

```
  1 - MEKMSHDPIA ADIGTQVSDN ALHGVTAGST ALTSVTGLVP AGADEVSAQA ATAFTSEGIQ
 61 - LLASNASAQD QLHRAGEAVQ DVARTYSQID DGAAGVFAE
```

>Rv3873: 368 aa - M. tuberculosis - SEQ.ID.No.5

```
  1 - MLWHAMPPEL NTARLMAGAG PAPMLAAAAG WQTLSAALDA QAVELTARLN SLGEAWTGGG
 61 - SDKALAAATP MVVWLQTAST QAKTRAMQAT AQAAAYTQAM ATTPSLPEIA ANHITQAVLT
121 - ATNFFGINTI PIALTEMDYF IRMWNQAALA MEVYQAETAV NTLFEKLEPM ASILDPGASQ
181 - STTNPIFGMP SPGSSTPVGQ LPPAATQTLG QLGEMSGPMQ QLTQPLQQVT SLFSQVGGTG
241 - GGNPADEEAA QMGLLGTSPL SNHPLAGGSG PSAGAGLLRA ESLPGAGGSL TRTPLMSQLI
301 - EKPVAPSVMP AAAAGSSATG GAAPVGAGAM GQGAQSGGST RPGLVAPAPL AQEREEDDED
361 - DWDEEDDW
```

>Rv3878: 280 aa - M. tuberculosis - SEQ.ID.No.3

```
  1 - MAEPLAVDPT GLSAAAAKLA GLVFPQPPAP IAVSGTDSVV AAINETMPSI ESLVSDGLPG
 61 - VKAALTRTAS NMNAAADVYA KTDQSLGTSL SQYAFGSSGE GLAGVASVGG QPSQATQLLS
121 - TPVSQVTTQL GETAAELAPR VVATVPQLVQ LAPHAVQMSQ NASPIAQTIS QTAQQAAQSA
181 - QGGSGPMPAQ LASAEKPATE QAEPVHEVTN DDQGDQGDVQ PAEVVAAARD EGAGASPGQQ
241 - PGGGVPAQAM DTGAGARPAA SPLAAPVDPS TPAPSTTTTL
```

>Rv3879c: 729 aa - M. tuberculosis - SEQ.ID.No.6

```
  1 - MSITRPTGSY ARQMLDPGGW VEADEDTFYD RAQEYSQVLQ RVTDVLDTCR QQKGHVFEGG
 61 - LWSGGAANAA NGALGANINQ LMTLQDYLAT VITWHRHIAG LIEQAKSDIG NNVDGAQREI
121 - DILENDPSLD ADERHTAINS LVTATHGANV SLVAETAERV LESKNWKPPK NALEDLLQQK
181 - SPPPPDVPTL VVPSPGTPGT PGTPITPGTP ITPGTPITPI PGAPVTPITP TPGTPVTPVT
241 - PGKPVTPVTP VKPGTPGEPT PITPVTPPVA PATPATPATP VTPAPAPHPQ PAPAPAPSPG
301 - PQPVTPATPG PSGPATPGTP GGEPAPHVKP AALAEQPGVP GQHAGGGTQS GPAHADESAA
361 - SVTPAAASGV PGARAAAAAP SGTAVGAGAR SSVGTAAASG AGSHAATGRA PVATSDKAAA
421 - PSTRAASART APPARPPSTD HIDKPDRSES ADDGTPVSMI PVSAARAARD AATAAASARQ
481 - RGRGDALRLA RRIAAALNAS DNNAGDYGFF WITAVTTDGS IVVANSYGLA YIPDGMELPN
541 - KVYLASADHA IPVDEIARCA TYPVLAVQAW AAFHDMTLRA VIGTAEQLAS SDPGVAKIVL
601 - EPDDIPESGK MTGRSRLEVV DPSAAAQLAD TTDQRLLDLL PPAPVDVNPP GDERHMLWFE
661 - LMKPMTSTAT GREAAHLRAF RAYAAHSQEI ALHQAHTATD AAVQRVAVAD WLYWQYVTGL
721 - LDRALAAAC
```

>Rv1979c: 481 aa - M. tuberculosis -

```
  1 - VGPRTRGYAI HKLGFCSVVM LGINSIIGAG IFLTPGEVIG LAGPFAPMAY VLAGIFAGVV
 61 - AIVFATAARY VRTNGASYAY TTAAFGRRIG IYVGVTHAIT ASIAWGVLAS FFVSTLLRVA
121 - FPDKAWADAE QLFSVKTLTF LGFIGVLLAI NLFGNRAIKW ANGTSTVGKA FALSAFIVGG
181 - LWIITTQHVN NYATAWSAYS ATPYSLLGVA EIGKGTFSSM ALATIVALYA FTGFESIANA
241 - AEEMDAPDRN LPRAIPIAIF SVGAIYLLTL TVAMLLGSNK IAASDDTVKL AAAIGNATFR
301 - IIIVVGALIS MFGINVAASF GAPRLWTALA DSGVLPTRLS RKNQYDVPMV SFAITASLAL
361 - AFPLALRFDN LHLTGLAVIA RFVQFIIVPI ALIALARSQA VEHAAVRRNA FTDKVLPLVA
421 - IVVSVGLAVS YDYRCIFLVR GGPNYFSIAL IVITFVVVPA MAYLHYYRII RRVGDRPSTR
```

>Rv1769: 414 aa - M. tuberculosis -

FIGURE 7 (CONTINUED)

```
  1 - VHEVAAREQR SDGPMRLDAQ GRLQRYEEAF ADYDAPFAFV DLDAMWGNAD QLLARAGDKP
 61 - IRVASKSLRC RPLQREILDA SERFDGLLTF TLTETLWLAG QGFSNLLLAY PPTDRAALRA
121 - LGELTAKDPD GAPIVMVDSV EHLDLIERTT DKPVRLCLDF DAGYWRAGGR IKIGSKRSPL
181 - HTPEQARALA VEIARRPALT LAALMCYEAH IAGLGDNVAG KRVHNAIIRR MQRMSFEELR
241 - ERRARAVELV REVADIKIVN AGGTGDLQLV AQEPLITEAT AGSGFYAPTL FDSYSTFTLQ
301 - PAAMFALPVC RRPGAKTVTA LGGGYLASGV GAKDRMPTPY LPVGLKLNAL EGTGEVQTPL
361 - SGDAARRLKL GDKVYFRHTK AGELCERFDH LHLVRGAEVV DTVPTYRGEG RTFL
```

FIGURE 8

983 1674 bp - M. tuberculosis - SEQ.ID.NO.11

```
                                                        ggtca
      gcgagttcggcggctagtcggtctacctcagggtctttg
      atattcagcgccacaggtagatggtaccagcaaatagcc
      actatctacctaacgcgtgctgtgccgtgcggtagctac
      tgaaaatccgagatgtcaaaggcagcgtctggatacgct
      gtatgcgcgcagggatggtgatcgaggcggaggggcggc
   1 - gtg tca ttt ctg gtc gtg gtt ccc gag ttc
  31 - ttg acg tcc gcg gca gcg gat gtg gag aac
  61 - ata ggt tcc aca ctg cgc gcg gcg aat gcc
  91 - gcg gct gcc gcc tcg acc acc gcg ctt gcg
  21 - gcc gct ggc gct gat gag gta tcg gcg gcg
  51 - gtg gca gcg ctg ttt gcc agg ttc ggt cag
  81 - gaa tat caa gcg gtc agc gcg cag gcg agc
  11 - gct ttc cat caa cag ttc gtg cag acg ctg
  41 - aac tcg gcg tca gga tcg tat gcg gcc gcg
  71 - gag gcc acc atc gcg tca cag ttg cag acc
  01 - gcg cag cac gat ctg ctg ggc gcg gtc aat
  31 - gca cca acc gaa acg ttg ttg ggg cgt ccg
  61 - cta atc ggc gac gga gca ccc ggg acg gca
  91 - acg agt ccg aat ggc ggg gcg ggt ggg ctg
  21 - ctg tac ggc aac ggc ggc aac ggt tat tcc
  51 - gcg acg gcg tcg ggg gtc ggc ggc ggg gcc
  81 - ggc ggt tcc gcg ggg ttg atc ggc aat ggc
  11 - ggc gcc ggg gga gcc ggc gga ccc aac gcc
  41 - ccc ggg gga gcc ggc ggc aac ggt ggc tgg
  71 - ctg ctc ggc aac ggc ggg atc ggc ggg ccc
  01 - ggg ggc gcg tcg agc atc ccc ggc atg agt
  31 - ggt gga gcc ggc gga acc ggc ggt gcc gca
  61 - gga ctt ttg ggc tgg gga gcg aac ggc gga
  91 - gcc ggc ggc ctc ggt gat gga gtc ggt gtc
  21 - gat cgt ggc acg ggc ggc gcc gga ggc cgc
  51 - ggc ggc ctg ttg tat ggc gga tac ggc gtc
  81 - agt ggg cca ggc ggc gac ggc aga acc gtc
  11 - ccg ctg gag ata att cat gtc aca gag ccg
  41 - acg gta cat gcc aac gtc aac ggc gga ccg
  71 - acg tca acc att ctg gtc gac acc gga tcc
  01 - gct ggt ctt gtt gtc tcg cct gag gat gtc
  31 - ggg gga atc ctg gga gtg ctt cac atg ggc
  61 - ctc cca acc gga ttg agc atc agc ggt tac
  91 - agc ggg ggg ctg tac tac atc ttc gcc acg
  21 - tat acc acg acg gtg gac ttc ggg aat ggc
  51 - atc gtc acc gcg ccg acc gcc gtt aat gtc
  81 - gtc ctc ttg tcc atc cca acg tcc ccc ttc
  11 - gcc att tcg acc tac ttc agc gcc ttg ctg
  41 - gcc gat ccg aca aca act ccg ttc gaa gcc
  71 - tat ttc ggt gcc gtc ggc gtg gac ggc gtt
  01 - ctg gga gtt ggg ccc aat gcg gtg gga cca
  31 - ggc ccc agc att ccg acg atg gcg tta ccg
  61 - ggt gac ctc aac cag gga gtg ctc atc gac
  91 - gca ccc gca ggt gag ctc gtg ttc ggt ccc
  21 - aac ccg cta cct gcg ccc aac gtc gag gtc
  51 - gtc gga tcg ccg atc acc acc ctg tac gta
  81 - aag atc gat ggt ggg act ccc ata ccc gtc
  11 - ccc tcg atc atc gat tcc ggt ggg gta acg
  41 - gga acc atc ccg tca tat gtc atc gga tcc
  71 - gga acc ctg ccg gcg aac aca aac att gag
  01 - gtc tac acc agc ccc ggc ggt gat cgg ctc
  31 - tac gcg ttc aac aca aac gat tac cgc ccg
  61 - acc gtc att tca tcc ggc ctg atg aat acc
  91 - ggg ttc ttg ccc ttc aga ttc cag ccg gtg
 221 - tac atc gac tac agc ccc agc ggt ata ggg
```

FIGURE 8 (CONTINUED)



651 - aca aca gtc ttt ga[redacted]t ccg gcg
tgatcgagcctgttcgcc[redacted]aatgtcgccgcctggctt
gtcatccccgactgaacatacgaaacatgcgccataata
ttgccgcctccggtgcatattggatcgtcgggagcacac
aagtttatggtcttagagctatacagcggaccgattgtc
ggcaacgacccgccgccccacaacatgctggagaaacca
ctgga

v1983: 1674 bp - M. tuberculosis -
gtcatttctggtcgtggttcccgagttcttgacgtccgcggcagcggatgtggagaac
aggttccacactgcgcgcggcgaatgccgcggctgccgcctcgaccaccgcgcttgcg
cgctggcgctgatgaggtatcggcggcggtggcagcgctgtttgccaggttcggtcag
atatcaagcggtcagcgcgcaggcgagcgctttccatcaacagttcgtgcagacgctg
gttcggcgtcaggatcgtatgcggccgcggaggccaccatcgcgtcacagttgcagacc
gcagcacgatctgctgggcgcggtcaatgcaccaaccgaaacgttgttggggcgtccg
aatcggcgacggagcacccgggacggcaacgagtccgaatggcggggcgggtgggctg
gtacggcaacggcggcaacggttattccgcgacggcgtcgggggtcggcggcggggcc
cggttccgcggggttgatcggcaatggcggcgccgggggagccggcggacccaacgcc
cggggggagccggcggcaacggtggctggctgctcggcaacggcgggatcggcgggccc
gggcgcgtcgagcatccccggcatgagtggtggagccggcggaaccggcggtgccgca
acttttgggctggggagcgaacggcggagccggcggcctcggtgatggagtcggtgtc
tcgtggcacgggcggcgccggaggccgcggcggcctgttgtatggcggatacggcgtc
ttgggccaggcggcgacggcagaaccgtcccgctggagataattcatgtcacagagccg
cggtacatgccaacgtcaacggcggaccgacgtcaaccattctggtcgacaccggatcc
ctggtcttgttgtctcgcctgaggatgtcgggggaatcctgggagtgcttcacatgggc
acccaaccggattgagcatcagcggttacagcgggggggctgtactacatcttcgccacg
ctaccacgacggtggacttcgggaatggcatcgtcaccgcgccgaccgccgttaatgtc
acctcttgtccatcccaacgtccccccttcgccatttcgacctacttcagcgccttgctg
ccgatccgacaacaactccgttcgaagcctatttcggtgccgtcggcgtggacggcgtt
cgggagttgggcccaatgcggtgggaccaggccccagcattccgacgatggcgttaccg
ctgacctcaaccagggagtgctcatcgacgcacccgcaggtgagctcgtgttcggtccc
acccgctacctgcgcccaacgtcgaggtcgtcggatcgccgatcaccaccctgtacgta
cgatcgatggtgggactcccatacccgtcccctcgatcatcgattccggtggggtaacg
caaccatcccgtcatatgtcatcggatccggaaccctgccggcgaacacaaacattgag
actacaccagccccggcggtgatcggctctacgcgttcaacacaaacgattaccgcccg
ccgtcatttcatccggcctgatgaataccgggttcttgcccttcagattccagccggtg
acatcgactacagccccagcggtatagggacaacagtctttgatcatccggcg

FIGURE 8 (CONTINUED)

SEQ. ID. NO. 8

1986: 597 bp - M. tuberculosis -

```
                                                    tgtag
                                                gcgctccgcggccgcatcgaagctgcccagttcgaccac
                                              ggcagccaatgcggccagctgtggaccgtcaagctgcgg
                                              atccaccatctcaggtgtagaccatctgcggagcgtcgc
                                              actgcacattaataatgctaatgtaaatgaagaattatt
                                              agctatactgacccatacaaactgcctagtgtcgattgc
1 - gtg aac tca cca ctg gtc gtc ggc ttc ctg
1 - gcc tgc ttc acg ctg atc gcc gcg att ggc
1 - gcg cag aac gca ttc gtg ctg cgg cag gga
1 - atc cag cgt gag cac gtg ctg ccg gtg gtg
1 - gcg ctg tgc acg gtg tcc gac atc gtg ctg
1 - atc gcc gcc ggt atc gcg ggg ttc ggc gca
1 - ttg atc ggc gca cat ccg cgt gcg ctc aat
1 - gtc gtc aag ttt ggc ggc gcc gcc ttc cta
1 - atc ggc tac ggg cta ctt gcg gcc cgg cgg
1 - gcg tgg cga cct gtt gcg ctg atc cca tct
1 - ggc gcc acg ccg gtt cgc tta gcc gag gtc
1 - ctg gtg acc tgt gcg gca ttc acg ttc ctc
1 - aac cca cac gtc tac ctc gac acc gtc gtg
1 - ttg cta ggc gcg ctg gcc aac gag cac agc
1 - gac cag cgc tgg ctg ttc ggc ctc ggc gcg
1 - gtc aca gcc agt gcg gta tgg ttc gcc acc
1 - ctc ggg ttc gga gcc ggc cgg ttg cgc ggg
1 - ctg ttc acc aac ccc ggc tcg tgg aga atc
1 - ctc gac ggc ctg atc gcg gtc atg atg gtt
1 - gcg ctg gga atc tcg ctg acc gtg acc
                                                tagtacagcacgtgtgcacacgcgggttggaccacgtga
                                                tcgtcgatgggcacataccgttcggcaggagggcgcgcg
                                                gtcagtctgcacaactcagtcaccagctgacacgccgac
                                                ggcggcctcgcccgggcgtgtcggcgccaccagtgcaca
                                                ttcggcgtgacgcggccctacggatcgtgttggagctgt
                                              agccc
```

1986: 597 bp - M. tuberculosis -

```
aactcaccactggtcgtcggcttcctggcctgcttcacgctgatcgccgcgattggc
cagaacgcattcgtgctgcggcagggaatccagcgtgagcacgtgctgccggtggtg
ctgtgcacggtgtccgacatcgtgctgatcgccgccggtatcgcggggttcggcgca
atcggcgcacatccgcgtgcgctcaatgtcgtcaagtttggcggcgccgccttccta
ggctacgggctacttgcggcccggcgggcgtggcgacctgttgcgctgatcccatct
gccacgccggttcgcttagccgaggtcctggtgacctgtgcggcattcacgttcctc
ccacacgtctacctcgacaccgtcgtgttgctaggcgcgctggccaacgagcacagc
cagcgctggctgttcggcctcggcgcggtcacagccagtgcggtatggttcgccacc
gggttcggagccggccggttgcgcgggctgttcaccaaccccggctcgtggagaatc
gacggcctgatcgcggtcatgatggttgcgctgggaatctcgctgaccgtgacc
```

FIGURE 8 (CONTINUED)

3872: 297 bp - M. tuberculosis - SEQ. ID. NO. 9

```
                                                    ggccc
      cctacatcgagcctccagaagaagtgttcgcagcacccc
      caagcgccggttaagattatttcattgccggtgtagcag
      gacccgagctcagcccggtaatcgagttcgggcaatgct
      gaccatcgggtttgtttccggctataaccgaacggtttg
      tgtacgggatacaaatacagggagggaagaagtaggcaa
  1 - atg gaa aaa atg tca cat gat ccg atc gct
 31 - gcc gac att ggc acg caa gtg agc gac aac
 61 - gct ctg cac ggc gtg acg gcc ggc tcg acg
 91 - gcg ctg acg tcg gtg acc ggg ctg gtt ccc
121 - gcg ggg gcc gat gag gtc tcc gcc caa gcg
151 - gcg acg gcg ttc aca tcg gag ggc atc caa
181 - ttg ctg gct tcc aat gca tcg gcc caa gac
  1 - cag ctc cac cgt gcg ggc gaa gcg gtc cag
211 - gac gtc gcc cgc acc tat tcg caa atc gac
271 - gac ggc gcc gcc ggc gtc ttc gcc gaa
      taggcccccaacacatcggagggagtgatcaccatgctg
      tggcacgcaatgccaccggagctaaataccgcacggctg
      atggccggcgcgggtccggctccaatgcttgcggcggcc
      gcgggatggcagacgctttcggcggctctggacgctcag
      gccgtcgagttgaccgcgcgcctgaactctctgggagaa
      gcctg
```

```
>3872: 297 bp - M. tuberculosis -
ggaaaaaatgtcacatgatccgatcgctgccgacattggcacgcaagtgagcgacaac
ctgcacggcgtgacggccggctcgacggcgctgacgtcggtgaccgggctggttccc
gggggccgatgaggtctccgcccaagcggcgacggcgttcacatcggagggcatccaa
gctggcttccaatgcatcggcccaagaccagctccaccgtgcgggcgaagcggtccag
cgtcgcccgcacctattcgcaaatcgacgacggcgccgccggcgtcttcgccgaa
```

FIGURE 8 (CONTINUED)

SEQ.ID.No.12

3873: 1104 bp - M. tuberculosis -

```
                                                    atgag
        gtctccgcccaagcggcgacggcgttcacatcggagggc
        atccaattgctggcttccaatgcatcggcccaagaccag
        ctccaccgtgcgggcgaagcggtccaggacgtcgcccgc
        acctattcgcaaatcgacgacggcgccgccggcgtcttc
        gccgaataggcccccaacacatcggagggagtgatcacc
  1 - atg ctg tgg cac gca atg cca ccg gag cta
 31 - aat acc gca cgg ctg atg gcc ggc gcg ggt
 61 - ccg gct cca atg ctt gcg gcg gcc gcg gga
 91 - tgg cag acg ctt tcg gcg gct ctg gac gct
121 - cag gcc gtc gag ttg acc gcg cgc ctg aac
151 - tct ctg gga gaa gcc tgg act gga ggt ggc
181 - agc gac aag gcg ctt gcg gct gca acg ccg
211 - atg gtg gtc tgg cta caa acc gcg tca aca
241 - cag gcc aag acc cgt gcg atg cag gcg acg
271 - gcg caa gcc gcg gca tac acc cag gcc atg
301 - gcc acg acg ccg tcg ctg ccg gag atc gcc
331 - gcc aac cac atc acc cag gcc gtc ctt acg
361 - gcc acc aac ttc ttc ggt atc aac acg atc
391 - ccg atc gcg ttg acc gag atg gat tat ttc
421 - atc cgt atg tgg aac cag gca gcc ctg gca
451 - atg gag gtc tac cag gcc gag acc gcg gtt
481 - aac acg ctt ttc gag aag ctc gag ccg atg
511 - gcg tcg atc ctt gat ccc ggc gcg agc cag
541 - agc acg acg aac ccg atc ttc gga atg ccc
571 - tcc cct ggc agc tca aca ccg gtt ggc cag
601 - ttg ccg ccg gcg gct acc cag acc ctc ggc
631 - caa ctg ggt gag atg agc ggc ccg atg cag
661 - cag ctg acc cag ccg ctg cag cag gtg acg
691 - tcg ttg ttc agc cag gtg ggc ggc acc ggc
721 - ggc ggc aac cca gcc gac gag gaa gcc gcg
751 - cag atg ggc ctg ctc ggc acc agt ccg ctg
781 - tcg aac cat ccg ctg gct ggt gga tca ggc
811 - ccc agc gcg ggc gcg ggc ctg ctg cgc gcg
841 - gag tcg cta cct ggc gca ggt ggg tcg ttg
871 - acc cgc acg ccg ctg atg tct cag ctg atc
901 - gaa aag ccg gtt gcc ccc tcg gtg atg ccg
931 - gcg gct gct gcc gga tcg tcg gcg acg ggt
961 - ggc gcc gct ccg gtg ggt gcg gga gcg atg
991 - ggc cag ggt gcg caa tcc ggc ggc tcc acc
1021 - agg ccg ggt ctg gtc gcg ccg gca ccg ctc
1051 - gcg cag gag cgt gaa gaa gac gac gag gac
1081 - gac tgg gac gaa gag gac gac tgg
          tgagctcccgtaatgacaacagacttcccggccacccgg
          gccggaagacttgccaacattttggcgaggaaggtaaag
          agagaaagtagtccagcatggcagagatgaagaccgatg
          ccgctaccctcgcgcaggaggcaggtaatttcgagcgga
          tctccggcgacctgaaaacccagatcgaccaggtggagt
          cgacg
```

Rv3873: 1104 bp - M. tuberculosis -

```
cgctgtggcacgcaatgccaccggagctaaataccgcacggctgatggccggcgcgggt
cggctccaatgcttgcggcggccgcgggatggcagacgctttcggcggctctggacgct
aggccgtcgagttgaccgcgcgcctgaactctctgggagaagcctggactggaggtggc
gcgacaaggcgcttgcggctgcaacgccgatggtggtctggctacaaaccgcgtcaaca
aggccaagacccgtgcgatgcaggcgacggcgcaagccgcggcatacacccaggccatg
ccacgacgccgtcgctgccggagatcgccgccaaccacatcacccaggccgtccttacg
ccaccaacttcttcggtatcaacacgatcccgatcgcgttgaccgagatggattatttc
tccgtatgtggaaccaggcagccctggcaatggaggtctaccaggccgagaccgcggtt
acacgctttcgagaagctcgagccgatggcgtcgatccttgatcccggcgcgagccag
gcacgacgaacccgatcttcggaatgccctcccctggcagctcaacaccggttggccag
```

cgccggcggctacccagaccctcgg[redacted]actgggtgagatgagcggcccgatgcag
tga[redacted]ccagccgctgcagcaggtgacgtcgttgttcagccaggtgggcggcaccggc
gc[redacted]ccagccgacgaggaagccgcgcagatgggcctgctcggcaccagtccgctg
accatccgctggctggtggatcaggccccagcgcgggcgcgggcctgctgcgcgcg
cgctacctggcgcaggtgggtcgttgacccgcacgccgctgatgtctcagctgatc
agccggttgccccctcggtgatgccggcggctgctgccggatcgtcggcgacgggt
ccgctccggtgggtgcgggagcgatgggccagggtgcgcaatccggcggctccacc
cgggtctggtcgcgccggcaccgctcgcgcaggagcgtgaagaagacgacgaggac
gggacgaagaggacgactgg

FIGURE 8 (CONTINUED)

SEQ. ID. NO. 10

3878: 840 bp - M. tuberculosis -

```
                                                        tgctg
      tggatcaccggggtgtacgacacggtccgcaatatccgg
      ttctgagccggatcggctgattggcggttcctgacagaa
      catcgaggacacggcgcaggtttgcataccttcggcgcc
      cgacaaattgctgcgattgagcgtgtggcgcgtccggta
      aaatttgctcgatggggaacacgtataggagatccggca
  1 - atg gct gaa ccg ttg gcc gtc gat ccc acc
 31 - ggc ttg agc gca gcg gcc gcg aaa ttg gcc
 61 - ggc ctc gtt ttt ccg cag cct ccg gcg ccg
 91 - atc gcg gtc agc gga acg gat tcg gtg gta
121 - gca gca atc aac gag acc atg cca agc atc
151 - gaa tcg ctg gtc agt gac ggg ctg ccc ggc
181 - gtg aaa gcc gcc ctg act cga aca gca tcc
211 - aac atg aac gcg gcg gcg gac gtc tat gcg
241 - aag acc gat cag tca ctg gga acc agt ttg
271 - agc cag tat gca ttc ggc tcg tcg ggc gaa
301 - ggc ctg gct ggc gtc gcc tcg gtc ggt ggt
331 - cag cca agt cag gct acc cag ctg ctg agc
361 - aca ccc gtg tca cag gtc acg acc cag ctc
391 - ggc gag acg gcc gct gag ctg gca ccc cgt
421 - gtt gtt gcg acg gtg ccg caa ctc gtt cag
451 - ctg gct ccg cac gcc gtt cag atg tcg caa
481 - aac gca tcc ccc atc gct cag acg atc agt
511 - caa acc gcc caa cag gcc gcc cag agc gcg
541 - cag ggc ggc agc ggc cca atg ccc gca cag
571 - ctt gcc agc gct gaa aaa ccg gcc acc gag
601 - caa gcg gag ccg gtc cac gaa gtg aca aac
631 - gac gat cag ggc gac cag ggc gac gtg cag
661 - ccg gcc gag gtc gtt gcc gcg gca cgt gac
691 - gaa ggc gcc ggc gca tca ccg ggc cag cag
721 - ccc ggc ggg ggc gtt ccc gcg caa gcc atg
751 - gat acc gga gcc ggt gcc cgc cca gcg gcg
781 - agt ccg ctg gcg gcc ccc gtc gat ccg tcg
811 - act ccg gca ccc tca aca acc aca acg ttg

      tagaccgggcctgccagcggctccgtctcgcacgcagcg
      cctgttgctgtcctggcctcgtcagcatgcggcggccag
      ggcccggtcgagcaacccggtgacgtattgccagtacag
      ccagtccgcgacggccacacgctggacggccgcgtcagt
      cgcagtgtgcgcttggtgcagggcaatctcctgtgagtg
     ggcag
```

```
>3878: 840 bp - M. tuberculosis -
ggctgaaccgttggccgtcgatcccaccggcttgagcgcagcggccgcgaaattggcc
cctcgtttttccgcagcctccggcgccgatcgcggtcagcggaacggattcggtggta
agcaatcaacgagaccatgccaagcatcgaatcgctggtcagtgacgggctgcccggc
gaaagccgccctgactcgaacagcatccaacatgaacgcggcggcggacgtctatgcg
gaccgatcagtcactgggaaccagtttgagccagtatgcattcggctcgtcgggcgaa
cctggctggcgtcgcctcggtcggtggtcagccaagtcaggctacccagctgctgagc
acccgtgtcacaggtcacgacccagctcggcgagacggccgctgagctggcaccccgt
tgttgcgacggtgccgcaactcgttcagctggctccgcacgccgttcagatgtcgcaa
cgcatcccccatcgctcagacgatcagtcaaaccgcccaacaggccgcccagagcgcg
gggcggcagcggcccaatgcccgcacagcttgccagcgctgaaaaaccggccaccgag
agcggagccggtccacgaagtgacaaacgacgatcagggcgaccagggcgacgtgcag
ggccgaggtcgttgccgcggcacgtgacgaaggcgccggcgcatcaccgggccagcag
cggcgggggcgttcccgcgcaagccatggataccggagccggtgcccgcccagcggcg
tccgctggcggccccccgtcgatccgtcgactccggcaccctcaacaaccacaacgttg
```

FIGURE 6 (CONTINUED)

3879c: 2187 bp - M. tuberc[redacted]sis - SEQ.ID.NO.13

```
                                                          cccgt
      gcgacagcgcgqccagcgcaagcgaggtgaccacccggc
      tgatcgcccaagcggtgcatcatgcgcgcggattcaacg
      ggttactgcgaataccggcgcggggtggatccagcggcc
      gaqccggcgtgaaatgccggaggccaaccggacggtgat
      ccgcgaggcgatctggcggtttggggagggcagtagggg
   1 - atg agt att acc agg ccg acg ggc agc tat
  31 - gcc aga cag atg ctg gat ccg ggc ggc tgg
  61 - gtg gaa gcc gat gaa gac act ttc tat gac
  91 - cgg gcc cag gaa tat agc cag gtt ttg caa
 121 - agg gtc acc gat gta ttg gac acc tgc cgc
 151 - cag cag aaa ggc cac gtc ttc gaa ggc ggc
 181 - cta tgg tcc ggc ggc gcc gcc aat gct gcc
 211 - aac ggc gcc ctg ggt gca aac atc aat caa
 241 - ttg atg acg ctg cag gat tat ctc gcc acg
 271 - gtg att acc tgg cac agg cat att gcc ggg
 301 - ttg att gaq caa gct aaa tcc gat atc ggc
 331 - aat aat gtg gat ggc gct caa cgg gag atc
 361 - gat atc ctg gag aat gac cct agc ctg gat
 391 - gct gat gag cgc cat acc gcc atc aat tca
 421 - ttg gtc acg gcg acg cat ggg gcc aat gtc
 451 - agt ctg gtc gcc gag acc gct gag cgg gtg
 481 - ctg gaa tcc aag aat tgg aaa cct ccg aag
 511 - aac gca ctc gag gat ttg ctt cag cag aag
 541 - tcg ccg cca ccc cca gac gtg cct acc ctg
 571 - gtc gtg cca tcc ccg ggc aca ccg ggc aca
 601 - ccg gga acc ccg atc acc ccg gga acc ccg
 631 - atc acc ccg gga acc cca atc aca ccc atc
 661 - ccg gga gcq ccg gta act ccg atc aca cca
 691 - acg ccc ggc act ccc gtc acg ccg gtg acc
 721 - ccg ggc aag ccg gtc acc ccg gtg acc ccg
 751 - gtc aaa ccg ggc aca cca ggc gag cca acc
 781 - ccg atc acg ccg gtc acc ccc ccg gtc gcc
 811 - ccg gcc aca ccg gca acc ccg gcc acg ccc
 841 - gtt acc cca gct ccc gct cca cac ccg cag
 871 - ccg gct ccg gca ccg gcg cca tcg cct ggg
 901 - ccc cag ccg gtt aca ccg gcc act ccc ggt
 931 - ccg tct ggt cca gca aca ccg ggc acc cca
 961 - ggg ggc gag ccg gcg ccg cac gtc aaa ccc
 991 - gcg gcg ttg gcg gag caa cct ggt gtg ccg
1021 - ggc cag cat gcg ggc ggg ggg acg cag tcg
1051 - ggg cct gcc cat gcg gac gaa tcc gcc gcg
1081 - tcg gtg acg ccg gct gcq gcq tcc ggt gtc
1111 - ccq ggc gca cgg gcq gcq gcc gcc gcq ccg
1141 - agc ggt acc gcc gtg gga gcq ggc gcg cgt
1171 - tcg agc gtg ggt acg gcc gcq gcc tcg ggc
1201 - gcg ggg tcg cat gct gcc act ggg cgg gcg
1231 - ccg gtg gct acc tcg gac aag gcg gcg gca
1261 - ccg agc acg cgg gcg gcc tcg gcq cgg acg
1291 - gca cct cct gcc cgc ccg ccg tcg acc gat
1321 - cac atc gac aaa ccc gat cgc agc gag tct
1351 - gca gat gac ggt acg ccg gtg tcg atg atc
1381 - ccg gtg tcg gcg gct cgg gcg gca cgc gac
1411 - gcc gcc act gca gct gcc agc gcc cgc cag
1441 - cgt ggc cgc ggt gat gcg ctg cgg ttg gcg
1471 - cga cgc atc gcg gcg gcg ctc aac gcq tcc
1501 - gac aac aac gcg ggc gac tac ggg ttc ttc
1531 - tgg atc acc gcg gtg acc acc gac ggt tcc
1561 - atc gtc gtg gcc aac agc tat ggg ctg gcc
1591 - tac ata ccc gac ggg atg gaa ttg ccg aat
1621 - aag gtg tac ttg gcc agc gcg gat cac gca
1651 - atc ccg gtt gac gaa atc gca cgc tgt gcc
1681 - acc tac ccg gtt ttg gcc gtg caa gcc tgg
1711 - gcg gct ttc cac gac atg acg ctg cgg gcg
```

FIGURE 8 (CONTINUED)

- gtg atc ggt acc gcg ga[illegible]g ttg gcc agt
- tcg gat ccc ggt gtg gcc [illegible]ag att gtg ctg
- [illegible]g cca gat gac att ccg gag agc ggc aaa
- [illegible]g acg ggc cgg tcg cgg ctg gag gtc gtc
- gac ccc tcg gcg gcg gct cag ctg gcc gac
- act acc gat cag cgt ttg ctc gac ttg ttg
- ccg ccg gcg ccg gtg gat gtc aat cca ccg
- ggc gat gag cgg cac atg ctg tgg ttc gag
- ctg atg aag ccc atg acc agc acc gct acc
- ggc cgc gag gcc gct cat ctg cgg gcg ttc
- cgg gcc tac gct gcc cac tca cag gag att
- gcc ctg cac caa gcg cac act gcg act gac
- gcg gcc gtc cag cgt gtg gcc gtc gcg gac
- tgg ctg tac tgg caa tac gtc acc ggg ttg
- ctc gac cgg gcc ctg gcc gcc gca tgc

tgacgaggccaggacagcaacaggcgctgcgtgcgagac
ggagccgctggcaggcccggtctacaacgttgtggttgt
tgagggtgccggagtcgacggatcgacgggggccgccag
cggactcgccgctgggcgggcaccggctccggtatccat
ggcttgcgcgggaacgccccccgccgggctgctggcccgg
tgatg

179c: 2187 bp - M. tuberculosis -

gtattaccaggccgacgggcagctatgccagacagatgctggatccgggcggctgg
aagccgatgaagacactttctatgaccgggcccaggaatatagccaggttttgcaa
tcaccgatgtattggacacctgccgccagcagaaaggccacgtcttcgaaggcggc
ggtccggcggcgccgccaatgctgccaacggcgccctgggtgcaaacatcaatcaa
tgacgctgcaggattatctcgccacggtgattacctggcacaggcatattgccggg
ttgagcaagctaaatccgatatcggcaataatgtggatggcgctcaacgggagatc
tcctggagaatgaccctagcctggatgctgatgagcgccataccgccatcaattca
tcacggcgacgcatggggccaatgtcagtctggtcgccgagaccgctgagcgggtg
aatccaagaattggaaacctccgaagaacgcactcgaggatttgcttcagcagaag
cgccacccccagacgtgcctaccctggtcgtgccatccccgggcacaccgggcaca
gaaccccgatcacccccgggaaccccgatcacccccgggaaccccaatcacacccatc
gagcgccggtaactccgatcacaccaacgccccggcactcccgtcacgccggtgacc
gcaagccggtcaccccggtgaccccggtcaaaccgggcacaccaggcgagccaacc
tcacgccggtcaccccccccggtcgccccggccacaccggcaaccccggccacgccc
ccccagctcccgctccacacccgcagccggctccggcaccggcgccatcgcctggg
agccggttacaccggccactcccggtccgtctggtccagcaacaccgggcacccca
gcgagccggcgccgcacgtcaaacccgcggcgtlggcggagcaacctggtgtgccg
agcatgcgggcgggggacgcagtcgggggcctgcccatgcggacgaatccgccgcg
tgacgccggctgcggcgtccggtgtcccgggcgcacgggcggcggccgccgccgcg
gtaccgccgtgggagcgggcgcgcgttcgagcgtgggtacggccgcggcctcgggc
ggtcgcatgctgccactgggcgggcgccggtggctacctcggacaaggcggcggca
gcacgcgggcggcctcggcgcggacggcacctcctgcccgcccgccgtcgaccgat
tcgacaaacccgatcgcagcgagtctgcagatgacggtacgccggtgtcgatgatc
tgtcggcggctcgggcggcacgcgacgccgccactgcagctgccagcgcccgccag
gccgcggtgatgcgctgcggttggcgcgacgcatcgcggcggcgctcaacgcgtcc
acaacgcgggcgactacgggttcttctggatcaccgcggtgaccaccgacggttcc
tcgtggccaacagctatgggctggcctacatacccgacgggatggaattgccgaat
tgtacttggccagcgcggatcacgcaatcccggttgacgaaattgcacgctgtgcc
acccggtttggccgtgcaagcctgggcggcttccacgacatgacgctgcgggcg
ttcggtaccgcggagcagttggccagttcggatcccggtgtggccaagattgtgctg
tcagatgacattccggagagcggcaaaatgacgggccggtcgcggctggaggtcgtc
cctcggcggcggctcagctggccgacactaccgatcagcgtttgctcgacttgttg
ccggcgccggtggatgtcaatccaccgggcgatgagcggcacatgctgtggttcgag
ttgaagcccatgaccagcaccgctaccggccgcgaggccgctcatctgcgggcgttc
gcctacgctgcccactcacaggagattgccctgcaccaagcgcacactgcgactgac
tccgtccagcgtgtggccgtcgcggactggctgtactggcaatacgtcaccgggttg
taccgggcccctggccgccgcatgc